

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 14:49:43 ; Search time 23.62 Seconds
(without alignments)
482.950 Million cell updates/sec

Title: US-09-554-547-16
Perfect score: 851
Sequence: 1 MKAFVLSLTSTALTNAR.....GSYGGSYTPNVRSAGAAA 154

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: A.GeneSeq_1101.*
- 2: /SID8/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
- 3: /SID8/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
- 4: /SID8/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
- 5: /SID8/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
- 6: /SID8/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
- 7: /SID8/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
- 8: /SID8/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
- 9: /SID8/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
- 10: /SID8/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
- 11: /SID8/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
- 12: /SID8/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
- 13: /SID8/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
- 14: /SID8/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
- 15: /SID8/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
- 16: /SID8/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
- 17: /SID8/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
- 18: /SID8/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
- 19: /SID8/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
- 20: /SID8/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
- 21: /SID8/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
- 22: /SID8/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
- 23: /SID8/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851	100.0	154	20	AAV13500
2	180	21.2	187	21	AAAG05210
3	180	21.2	308	21	AAAG05209
4	180	21.2	358	21	AAAG05208
5	175.5	20.6	187	22	AAAB83953
6	175.5	20.6	263	21	AAAG36620
7	175.5	20.6	273	21	AAAG36619
8	175.5	20.6	309	21	AAAG36618
9	175.5	20.6	378	22	AAAM39500
10	175.5	20.6	389	22	AAAM41286
11	175	20.6	154	21	AAAG31995

12	175	20.6	256	22	AAAB30807
13	172	20.2	85	21	AAAG25162
14	172	20.2	130	21	AAAG25161
15	172	20.2	615	20	AAV40099
16	172	20.2	831	16	AAAB0168
17	171	20.1	110	21	AAAG32939
18	171	20.1	126	21	AAAB19197
19	170	20.0	136	21	AAAB19196
20	169.5	19.9	126	21	AAAG31996
21	167.5	19.7	97	21	AAAG12064
22	167.5	19.7	126	21	AAAG12063
23	166.5	19.6	147	21	AAAG21911
24	166.5	19.6	147	21	AAAG25740
25	166.5	19.6	157	21	AAAG20292
26	166	19.5	128	21	AAAB58406
27	165	19.4	175	21	AAAG24564
28	164.5	19.3	114	21	AAAG24565
29	164.5	19.3	114	21	AAAG24993
30	164	19.3	105	21	AAAG14030
31	164	19.3	147	21	AAAG14029
32	163.5	19.2	135	21	AAAG28869
33	163	19.2	132	21	AAAB19199
34	163	19.2	212	22	AAAG75147
35	163	19.2	218	21	AAAB57093
36	163	19.2	341	18	AAAB26553
37	163	19.2	341	22	AAAB81935
38	163	19.2	353	18	AAAB26552
39	163	19.2	353	19	AAAM50921
40	163	19.2	353	19	AAAM54362
41	162.5	19.1	107	21	AAAG24566
42	162.5	19.1	107	21	AAAG24594
43	161.5	19.0	133	21	AAAG42333
44	161.5	19.0	203	21	AAAG29851
45	161.5	19.0	203	22	AAAB67835

ALIGNMENTS

RESULT 1	
ID	AAV13500 standard; Protein: 154 AA.
XX	AAV13500;
XX	30-JUL-1999 (first entry)
XX	
DE	Tissue cement protein encoded by clone 64.
XX	
KW	Tissue cement protein; TCP; blood-feeding; ectoparasite; therapy;
KW	pharmaceutical; vaccine; tissue bonding; surgical wound; laceration;
KW	skin grafting; immunogen; infection; arthropod parasite.
OS	Rhipicephalus appendiculatus.
XX	
PN	WO9924567-A1.
XX	
PD	20-MAY-1999.
XX	
PF	12-NOV-1998; 98WO-GB03397.
XX	
PR	12-NOV-1997; 97GB-0023945.
XX	
PA	(NAEN-) NAT ENVIRONMENTAL RES COUNCIL.
XX	
PI	Nuttall Patricia A, Paesen GC;
XX	
DR	WPI; 1999-327399/27.
XX	
DR	N-PSDB; AAX55699.
XX	
PT	Tissue cement proteins produced by blood-feeding ectoparasites and
XX	related polynucleotides

Claim 1, Fig 7; 65pp; English.

The invention relates to tissue cement proteins (TCPs) produced by blood-feeding ectoparasites. The TCP, in a pharmaceutical composition, is useful for therapy, as a vaccine or vaccine component. The TCP itself is used to immunize an animal for production of such a vaccine. The TCPS are useful for bonding animal tissues. This may be temporary or permanent bonding and used in, e.g., repair of incised surgical wounds, lacerations, skin grafting, etc. The TCP can also be used as a protective immunogen in the control of diseases caused by infections transmitted by arthropod parasites. The tick TCPS provide a non-immunogenic tissue cement capable of bonding mammalian tissue with great strength. The hardening or elastic properties of the TCPS can be tailored for particular requirements.

Sequence 154 AA;

Query Match 100.0%; Score 851; DB 20; Length 154;
Best Local Similarity 100.0%; Pred. NO. 3,1e-68;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKAFVLSLSTAAALTNARAGRLGSDUDTFGRVHGMLYAGIERACPRGYPGLTASIGGE 60
1 mkafvllslstaaaltnaaragrllgsddltgrvhnlyagieragprgyplgltaslgge 60

61 VGARLGGNAGVGVSSYGYGRPSWGYPRGCGYGGYGGYGGYDQGFSAVAGYPRGYGY 120
61 vgarlggnagvgvssygygpswgyprgcygygygygygydggfsgavygyprgygyy 120

121 YPSGYGGYGGSGYSGSYTYTPNVRASAGAA 154
121 ypsygygygygygygygytytpnvrasagaa 154

RESULT 2
AAG05210
ID AAG05210 standard; Protein: 187 AA.

AAG05210;
17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 1529.

Arabidopsis thaliana.
EP1033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.
23-MAR-1999; 99US-0125788.
25-MAR-1999; 99US-0126284.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.
06-APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.
16-APR-1999; 99US-0129845.
19-APR-1999; 99US-0130077.
21-APR-1999; 99US-0130449.
23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891.
28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0132048.
30-APR-1999; 99US-0132407.

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

PR	04-MAY-1999	9905-01332484
PR	03-MAY-1999	9905-01332485
PR	06-MAY-1999	9905-01332486
PR	06-MAY-1999	9905-01332487
PR	07-MAY-1999	9905-01332488
PR	11-MAY-1999	9905-01334256
PR	14-MAY-1999	9905-01334257
PR	14-MAY-1999	9905-01334258
PR	14-MAY-1999	9905-01334259
PR	14-MAY-1999	9905-01334260
PR	18-MAY-1999	9905-01334768
PR	19-MAY-1999	9905-01334769
PR	20-MAY-1999	9905-01335124
PR	21-MAY-1999	9905-01335523
PR	24-MAY-1999	9905-01356539
PR	25-MAY-1999	9905-01366921
PR	27-MAY-1999	9905-01366922
PR	28-MAY-1999	9905-01367882
PR	01-JUN-1999	9905-01373722
PR	03-JUN-1999	9905-01373728
PR	04-JUN-1999	9905-01373752
PR	07-JUN-1999	9905-01377240
PR	08-JUN-1999	9905-01380694
PR	10-JUN-1999	9905-01386540
PR	10-JUN-1999	9905-01388447
PR	14-JUN-1999	9905-01391119
PR	16-JUN-1999	9905-01394512
PR	16-JUN-1999	9905-01394523
PR	17-JUN-1999	9905-01394532
PR	18-JUN-1999	9905-01394554
PR	18-JUN-1999	9905-01394555
PR	18-JUN-1999	9905-01394565
PR	18-JUN-1999	9905-01394577
PR	18-JUN-1999	9905-01394587
PR	18-JUN-1999	9905-01394597
PR	18-JUN-1999	9905-01394600
PR	18-JUN-1999	9905-01394611
PR	18-JUN-1999	9905-01394632
PR	18-JUN-1999	9905-01394646
PR	24-JUN-1999	9905-01400655
PR	24-JUN-1999	9905-01400654
PR	28-JUN-1999	9905-01400823
PR	29-JUN-1999	9905-01400891
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-01418412
PR	01-JUL-1999	9905-01418414
PR	02-JUL-1999	9905-01420055
PR	06-JUL-1999	9905-01423900
PR	08-JUL-1999	9905-01428073
PR	09-JUL-1999	9905-01429320
PR	12-JUL-1999	9905-01429377
PR	13-JUL-1999	9905-01435342
PR	14-JUL-1999	9905-01436324
PR	15-JUL-1999	9905-01440005
PR	16-JUL-1999	9905-01440085
PR	16-JUL-1999	9905-01440086
PR	19-JUL-1999	9905-01443311
PR	19-JUL-1999	9905-01443321
PR	20-JUL-1999	9905-01443532
PR	20-JUL-1999	9905-01446332
PR	20-JUL-1999	9905-01446884
PR	21-JUL-1999	9905-01446114
PR	21-JUL-1999	9905-01450086
PR	21-JUL-1999	9905-01450088

PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145102;
PR	23-JUL-1999;	9905-0145115;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145951;
PR	28-AUG-1999;	9905-0145386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148341;
PR	12-AUG-1999;	9905-0148349;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0148368;
PR	23-AUG-1999;	9905-0148902;
PR	23-AUG-1999;	9905-0149302;
PR	26-AUG-1999;	9905-0150584;
PR	26-AUG-1999;	9905-0150864;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0153363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	24-SEP-1999;	9905-0155569;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156599;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158363;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0158294;
PR	14-OCT-1999;	9905-0155329;
PR	14-OCT-1999;	9905-0155330;
PR	14-OCT-1999;	9905-0155637;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160767;

[illegible]

PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 14-JUN-1999; 990S-0138847.
PR 16-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 17-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 02-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.

PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149923.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151030.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.

PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 11-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 42.4%; Score 180; DB 21; Length 358;
 Matches 56; Conservative 4; Mismatches 40; Indels 32; Gaps 10;

QY 33 RVHG---NLVACIERAGPRGYPGLTASIGCEVGARLCGRACGVSSYGYGY---PSMCGY- 85
 Db 233 rshgpgasryag-----gyygrgsvpelfg--gymny99gslyg-yrnepplgys 281
 QY 86 ----PVC-GYC-----GYGCGYGYGYPDOGRGSAYGV--PCYVGYVYPPSGYGGY---CG 131
 Db 282 srfgpygsfgrgegygrgegaylgyprggygy99p9gy99g99gsyegag9 341
 QY 132 SYGSGYGSYTY 143
 Db 342 pygrgyssary 353

RESULT 5

AAB83953 standard; Protein: 187 AA.

AC AAB83953;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of the corn silk C3 gene.

KW Corn silk gene; C3 gene; C3 regulatory region; female sterile plant;

KW pathogen resistance; silk tissue; pistil tissue; stigma.

XX Zea mays.

OS CA2321269-A1.

PN 05-APR-2001.

PF 05-OCT-2000; 2000CA-2321269.

PR 05-OCT-1999; 99US-0412704.

PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

PI Tao T, Singh J, Oneillet T;

DR WPI: 2001-329438/35.

DR N-PSDB; AAF89925.

PT New regulatory regions of corn silk for directing expression of a gene

PS of interest in stigma of both monocot and dicot plants

XX Example 3; Page 49-50; 50pp; English.

CC The present sequence is encoded by a corn silk gene, the C3 gene. The

CC regulatory region of the C3 gene is useful for producing a female sterile
CC plant, or a plant with or improved pathogen resistance. A vector
CC comprising a chimeric gene in association with the regulatory region is
CC useful for producing a plant expressing a gene of interest within silk or
CC plastid tissue. The C3 regulatory region may be used to direct the
CC expression of a gene of interest in the stigma of both monocot and dicot
CC plants. Furthermore, the activity of the regulatory region may be
CC modified by the presence of other regulatory regions, for example,
CC enhancers, core promoter elements etc..

XX Sequence 187 AA:

Query Match 20.6% Score 175.5; DB 22; Length 187;
Best Local Similarity 38.7% Pred. No. 2.4e-08;
Matches 55; Conservative 2; Mismatches 46; Indels 39; Gaps 7;

OY 47 PRCYVGL-----TASTCGEVGARLGGAGVSSSYG-----YPSWG-----YEP 87
DB 44 papepjldekewfgyrkhhggygnnpqyggsgnspqyggsgnspqygggykrhhpg 103
OY 88 GGY-----CGYCGYGGYGY--DQFGSAVCGYPCYGYTYPSGCGGCGSY 133
DB 104 ggygsggpggcggyggygsgpdygddnggsgtggngnag-----gyggyggyggy 158
OY 134 GSYG-GSYTPNVRAACAA 154
DB 159 ggygsgsgtappgyghggyga 180

RESULT 6

AAC36620 ID AAC36620 standard; Protein: 263 AA.

AC AAC36620;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44904.

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.

PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139453.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142877.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.

PR	23-JUL-1999	99US-0145518
PR	23-JUL-1999	99US-0145518
PR	23-JUL-1999	99US-01455276
PR	26-JUL-1999	99US-0145513
PR	27-JUL-1999	99US-0145518
PR	27-JUL-1999	99US-0145519
PR	28-JUL-1999	99US-0145521
PR	02-AUG-1999	99US-0146386
PR	02-AUG-1999	99US-0146388
PR	02-AUG-1999	99US-0146388
PR	06-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147416
PR	09-AUG-1999	99US-0147493
PR	09-AUG-1999	99US-0147535
PR	10-AUG-1999	99US-0148171
PR	11-AUG-1999	99US-0148319
PR	12-AUG-1999	99US-0148319
PR	13-AUG-1999	99US-0148365
PR	13-AUG-1999	99US-0148684
PR	16-AUG-1999	99US-0149368
PR	17-AUG-1999	99US-0149175
PR	18-AUG-1999	99US-0149426
PR	20-AUG-1999	99US-0149722
PR	20-AUG-1999	99US-0149723
PR	20-AUG-1999	99US-0149929
PR	22-AUG-1999	99US-0149930
PR	23-AUG-1999	99US-0149930
PR	25-AUG-1999	99US-0150566
PR	26-AUG-1999	99US-0150884
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151080
PR	30-AUG-1999	99US-0151103
PR	31-AUG-1999	99US-0151138
PR	01-SEP-1999	99US-0151330
PR	07-SEP-1999	99US-0152363
PR	10-SEP-1999	99US-0153070
PR	11-SEP-1999	99US-0153758
PR	15-SEP-1999	99US-0154018
PR	16-SEP-1999	99US-0154739
PR	22-SEP-1999	99US-0155479
PR	23-SEP-1999	99US-0155183
PR	24-SEP-1999	99US-0155659
PR	28-SEP-1999	99US-0156458
PR	29-SEP-1999	99US-0156596
PR	04-OCT-1999	99US-0157117
PR	04-OCT-1999	99US-0157753
PR	06-OCT-1999	99US-0158765
PR	07-OCT-1999	99US-0158029
PR	08-OCT-1999	99US-0158323
PR	12-OCT-1999	99US-0158369
PR	13-OCT-1999	99US-0159293
PR	13-OCT-1999	99US-0159294
PR	13-OCT-1999	99US-0159295
PR	14-OCT-1999	99US-0159329
PR	14-OCT-1999	99US-0159330
PR	14-OCT-1999	99US-0159637
PR	14-OCT-1999	99US-0159638
PR	18-OCT-1999	99US-0159584
PR	21-OCT-1999	99US-0160741
PR	21-OCT-1999	99US-0160767
PR	21-OCT-1999	99US-0160768
PR	21-OCT-1999	99US-0160770
PR	21-OCT-1999	99US-0160815
PR	21-OCT-1999	99US-0160815

Query Match	20.6%	Score 175.5	DB 21	Length 263	
Best Local Similarity	36.7%	Pred. No. 3,3e+08			
Matches 54	Conservative	9	Mismatches 43	Indels 41	Gaps 6
Pr 22-OCT-1999	99US-0160980				
Pr 22-OCT-1999	99US-0160981				
Pr 22-OCT-1999	99US-0160989				
Pr 23-OCT-1999	99US-0161404				
Pr 25-OCT-1999	99US-0161405				
Pr 25-OCT-1999	99US-0161406				
Pr 26-OCT-1999	99US-0161359				
Pr 26-OCT-1999	99US-0161360				
Pr 26-OCT-1999	99US-0161361				
Pr 28-OCT-1999	99US-0161920				
Pr 28-OCT-1999	99US-016192				
Pr 28-OCT-1999	99US-0161993				
Pr 29-OCT-1999	99US-0162142				
Query Match	20.6% <td>Score 175.5<td>DB 21<td>Length 263</td></td></td>	Score 175.5 <td>DB 21<td>Length 263</td></td>	DB 21 <td>Length 263</td>	Length 263	
Best Local Similarity	36.7% <td>Pred. No. 3,3e+08<td></td><td></td></td>	Pred. No. 3,3e+08 <td></td> <td></td>			
Matches 54 <td>Conservative<td>9<td>Mismatches 43<td>Indels 41<td>Gaps 6</td></td></td></td></td>	Conservative <td>9<td>Mismatches 43<td>Indels 41<td>Gaps 6</td></td></td></td>	9 <td>Mismatches 43<td>Indels 41<td>Gaps 6</td></td></td>	Mismatches 43 <td>Indels 41<td>Gaps 6</td></td>	Indels 41 <td>Gaps 6</td>	Gaps 6
Pr 22-OCT-1999	99US-0160980				
Pr 22-OCT-1999	99US-0160981				
Pr 22-OCT-1999	99US-0160989				
Pr 23-OCT-1999	99US-0161404				
Pr 25-OCT-1999	99US-0161405				
Pr 25-OCT-1999	99US-0161406				
Pr 26-OCT-1999	99US-0161359				
Pr 26-OCT-1999	99US-0161360				
Pr 26-OCT-1999	99US-0161361				
Pr 28-OCT-1999	99US-0161920				
Pr 28-OCT-1999	99US-016192				
Pr 28-OCT-1999	99US-0161993				
Pr 29-OCT-1999	99US-0162142				
Pr 22-OCT-1999	99US-0160980				
Pr 22-OCT-1999	99US-0160981				
Pr 22-OCT-1999	99US-0160989				
Pr 23-OCT-1999	99US-0161404				
Pr 25-OCT-1999	99US-0161405				
Pr 25-OCT-1999	99US-0161406				
Pr 26-OCT-1999	99US-0161359				
Pr 26-OCT-1999	99US-0161360				
Pr 26-OCT-1999	99US-0161361				
Pr 28-OCT-1999	99US-0161920				
Pr 28-OCT-1999	99US-016192				
Pr 28-OCT-1999	99US-0161993				
Pr 29-OCT-1999	99US-0162142				
Pr 22-OCT-1999	99US-0160980				
Pr 22-OCT-1999	99US-0160981				
Pr 22-OCT-1999	99US-0160989				
Pr 23-OCT-1999	99US-0161404				
Pr 25-OCT-1999	99US-0161405				
Pr 25-OCT-1999	99US-0161406				
Pr 26-OCT-1999	99US-0161359				
Pr 26-OCT-1999	99US-0161360				
Pr 26-OCT-1999	99US-0161361				
Pr 28-OCT-1999	99US-0161920				
Pr 28-OCT-1999	99US-016192				
Pr 28-OCT-1999	99US-0161993				
Pr 29-OCT-1999	99US-0162142				
Pr 22-OCT-1999	99US-0160980				
Pr 22-OCT-1999	99US-0160981				
Pr 22-OCT-1999	99US-0160989				
Pr 23-OCT-1999	99US-0161404				
Pr 25-OCT-1999	99US-0161405				
Pr 25-OCT-1999	99US-0161406				
Pr 26-OCT-1999	99US-0161359				
Pr 26-OCT-1999	99US-0161360				
Pr 26-OCT-1999	99US-0161361				
Pr 28-OCT-1999	99US-0161920				
Pr 28-OCT-1999	99US-016192				
Pr 28-OCT-1999	99US-0161993				
Pr 29-OCT-1999	99US-0162142				
Pr 22-OCT-1999	99US-0160980				
Pr 22-OCT-1999	99US-0160981				
Pr 22-OCT-1999	99US-0160989				
Pr 23-OCT-1999	99US-0161404				
Pr 25-OCT-1999	99US-0161405				

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PN	EP1033405-A2.		PR	01-JUL-1999;	99US-0142154.
XX			PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.		PR	06-JUL-1999;	99US-0142390.
XX			PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000EP-0301439.		PR	09-JUL-1999;	99US-0142820.
XX			PR	12-JUL-1999;	99US-0142877.
XX			PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	99US-0121825.	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	99US-0123180.	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	99US-0123548.	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	99US-0125788.	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	99US-0126264.	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	99US-0126785.	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	99US-0127462.	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	99US-0128234.	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	99US-0128714.	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	99US-0129845.	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	99US-0130077.	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144684.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149375.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139819.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151348.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.

PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137702.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145518.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145591.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.28; Score 172; DB 21; Length 85;
Best Local Similarity 45.58; Pred. No. 2.2e-08;
Matches 45; Conservative 5; Mismatches 23; Indels 26; Gaps 7;

Qy 43 ERAGPRGPGITASTIGEVGRLGGRAGVSVSYGYGYPMSGYRYGCGYGYG--GYG 100
Db 8 elvqpegyh-----ggh-----gghggyg--hyggyg---ghghgghngggyghldgyg 50
Qy 101 GYDQFGSAVAGYGPYGYGYRPSGCGYGSYGSYGG 139
Db 51 g---ghghghygggyggyg-----ggyggggygggygsggg 82

RESULT 14
AAG25161
ID AAG25161 standard; Protein; 130 AA.
XX
XX AAG25161;

XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 29112.
XX
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.
OS

XX EPI033405-A2.
PN

XX 06-SEP-2000.
PD

XX 25-FEB-2000; 2000EP-0301439.
PF

XX 25-FEB-1999; 99US-0121825.
PR

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128274.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 14:50:08 ; Search time 12.49 Seconds

(without alignments)
277.463 Million cell updates/sec

Title: US-09-554-547-16

Perfect score: 851
Sequence: 1 MKAFFVLLSTALTAALTNAR.....GSYGYSTYPNVASAGAA 154Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	20.2	832	1	US-08-209-747-2
2	172	20.2	832	1	US-08-458-298-2
3	171	20.1	126	1	US-08-963-168C-7
4	170	20.0	136	3	US-08-963-168C-6
5	163	19.2	132	3	US-08-963-168C-9
6	163	19.2	341	2	US-08-538-711A-8
7	163	19.2	341	4	US-08-725-027-8
8	163	19.2	353	4	US-08-538-711A-7
9	163	19.2	353	4	US-08-725-027-7
10	160.5	18.9	131	3	US-08-963-168C-8
11	154.5	18.2	745	2	US-09-010-928B-28
12	154.5	18.2	870	2	US-09-010-928B-2
13	153.5	18.0	116	3	US-08-963-168C-13
14	152.5	17.9	907	2	US-09-010-928B-4
15	151.5	17.8	106	3	US-08-963-168C-14
16	150.5	17.7	112	3	US-08-963-168C-16
17	146	17.2	604	4	US-08-556-978B-63
18	146	17.2	606	4	US-08-556-978B-23
19	146	17.2	606	4	US-09-247-806-6
20	146	17.2	606	4	US-09-247-806-8
21	144	16.9	111	3	US-08-963-168C-15
22	144	16.9	651	4	US-08-556-978B-19
23	144	16.9	651	4	US-09-247-806-1
24	144	16.9	718	1	US-08-425-069-2
25	144	16.9	718	2	US-08-317-844B-2
26	144	16.9	747	3	US-09-034-177-3
27	142.5	16.7	606	4	US-08-556-978B-21

28	142.5	16.7	606	4	US-09-247-806-4	Sequence 4, Appl
29	141.5	16.6	251	1	US-08-209-747-8	Sequence 8, Appl
30	141.5	16.6	251	1	US-08-458-298-8	Sequence 8, Appl
31	139.5	16.4	738	3	US-08-864-038A-3	Sequence 3, Appl
32	138	16.2	235	2	US-08-529-190B-1	Sequence 1, Appl
33	136.5	16.0	318	4	US-09-060-756-727	Sequence 127, App
34	136	16.0	334	4	US-09-060-756-728	Sequence 728, App
35	136	16.0	714	4	US-08-556-978B-61	Sequence 61, Appl
36	136	16.0	714	4	US-09-247-806-10	Sequence 10, Appl
37	134	15.7	1160	3	US-08-808-599A-24	Sequence 24, Appl
38	132	15.5	493	4	US-08-556-978B-59	Sequence 59, Appl
39	132	15.5	529	4	US-09-247-806-2	Sequence 2, Appl
40	128.5	15.1	595	1	US-08-425-069-4	Sequence 4, Appl
41	128.5	15.1	595	2	US-08-317-844B-4	Sequence 15, Appl
42	127.5	15.0	155	1	US-08-209-747-15	Sequence 15, Appl
43	127.5	15.0	155	1	US-08-458-298-15	Sequence 15, Appl
44	127	14.9	201	4	US-09-052-995-1	Sequence 1, Appl
45	127	14.9	201	4	US-09-053-003-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-209-747-2
; Sequence 2, Application US/08209747
; Patent No. 5733771
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Colgin, Mark
; TITLE OR INVENTION: CDNAS Encoding Minor Ampullate Spider
; TITLE OF INVENTION: Silk Proteins
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,747
; FILING DATE: 14-MAR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ. ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: N. clavipes
; TISSUE TYPE: minor ampullate gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..309
; US-08-209-747-2


```
Db 1 MNAVYICLSACLIATAVSAAGYGGAGS-MGCTGCMGGMNAG-----GFGCMGGMGCG 52
OY 61 VGARLGRAGVSSYGYGPPSMGYPYGGYGGYGGYDQGRGSAVGYGPGYGY 120
Db 53 KCG-FGCMGCGFPGCMGCGM-----GGP-GGFGMCGFPGCMGCGKCGCGMCGSGMGGFGM- 105
OY 121 YPSGYGGYGGYGC-GSYGC 139
Db 106 --GGGNAGFGCMGCGCGGCGFG 123

RESULT 4
US-08-963-168C-6
; Sequence 6, Application US/08963168C
; Patent No. 6127166
; GENERAL INFORMATION:
; APPLICANT: Bayley, Hagan
; APPLICANT: Cao, Qunping
; APPLICANT: Wang, Yunjaun
; TITLE OF INVENTION: MOULUSCAN LIGAMENT POLYPEPTIDES
; TITLE OF INVENTION: AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963.168C
; FILING DATE: 03-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; OTHER INFORMATION: This translation is for SEQ ID NOS:1 & 2.
; US-08-963-168C-6

Query Match 20.0%; Score 170; DB 3; Length 136;
Best Local Similarity 36.2%; Pred. No. 7.8e-09;
Matches 54; Conservative 17; Mismatches 52; Indels 26; Gaps 9;
```

```
RESULT 5
US-08-963-168C-9
; Sequence 9, Application US/08963168C
; Patent No. 6127166
; GENERAL INFORMATION:
; APPLICANT: Bayley, Hagan
; APPLICANT: Cao, Qunping
; APPLICANT: Wang, Yunjaun
; TITLE OF INVENTION: MOULUSCAN LIGAMENT POLYPEPTIDES
; TITLE OF INVENTION: AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963.168C
; FILING DATE: 03-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-963-168C-9

Query Match 19.2%; Score 163; DB 3; Length 132;
Best Local Similarity 35.4%; Pred. No. 3.3e-08;
Matches 52; Conservative 18; Mismatches 51; Indels 26; Gaps 8;

Db 1 MNAFYILSLSTALTNARAGRLGSDLPFGRVHGNLYAGIERAPRGYPTLSTIGE 60
OY 61 VGARLGRAGVSSYGYGPPSMGYPYGGYGGYGGYDQGRGSAVGYGPGYGY 117
Db 49 MCGGKGGFPGMCGGPGGFG--GIGGSGGCGMCGFPGMCGKCGCGMSSNGRGGMG 106
OY 118 GYYPGSGYGG-----GYGSGSYGSGYGC 139
Db 107 G--GNAGFGCMGCGMCGGCGSG--FGG 129

RESULT 6
US-08-538-711A-8
; Sequence 8, Application US/08538711A
; Patent No. 5994062
; GENERAL INFORMATION:
; APPLICANT: MUSHINE, JAMES, L.
; TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
```

```

ADDRESS: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154
```

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/538,711A
FILING DATE: 02-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4201
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEX: (212) 751-6849
INFORMATION FOR SEQ ID NO.: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 341
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Linear
MOLECULE TYPE: peptide

US-08-538-711A-8

Query Match 19.2%; Score 163; DB 2; Length 341;
Best Local Similarity 33.5%; Pred. No. 8.8e-08;

Matches 54; Conservative 11; Mismatches 56; Indels 40; Gaps

Dy 18 AARAGRLGSDLDTFGRVHGNLYAGIERAGPRGY-PGLTASI-GGEYGARLGGAGVGSS 75
 ::::| | | | | : | | | | |
Db 186 SRSRCR-----GSNFEGDSRRGGCGMFQPGPSNFRGSGDG--YSGRGRTGDGY 232

Oy 76 YGYCYPSMGICPYGCIGGY-GGYGGYGCTDQGFSAYGVTPGYTY---- 120
 |||| | : || | | | | | | | | | | | | | |
Db 233 NGYGGGPGGCNFGGSPGYGGGGRGYPGGNGGGGGYGDDNYGGGNNYNDFG 292

Oy 121 ----YPSGYG---GGYGS--YGSYSGGSYYPIPVVRASAG 151
 || || | : || | | | | | | | | | |
Db 293 NYNQQPNTYGMKSGFSGSRMRGEPYGGNGYPCGSGS 333

RESULT 7
US-08-725-027-8

Sequence 8, Application US/08725027
Patent No. 6251586

GENERAL INFORMATION:
APPLICANT: MULSHINE, JAMES, L.
APPLICANT: TOCKMAN, MELVIN, S.
TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
TITLE OF INVENTION: DNA THEREOF FOR USE IN EARLY CANCER DETECTION
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725.027
FILING DATE: 02-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/538,711
FILING DATE: 02-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4201US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 341
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-725-027-8

Query Match          19.2% Score 163; DB 4; Length 341;
Best Local Similarity 33.5%; Pred. No. 8.8e-08;
Matches 54; Conservative 11; Mismatches 56; Indels 40; Gaps

QY 18 AARAGRLCSDDLTDFGRVHGNLYAGIERAGPRGY-PGLTASI-GGEVGARLGRAGVSS 75
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 186 SRSR-----GCFGFGSRGGGNGFPGPSNFSGSNG--YGSGRGFGDY 232
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 YGYGYSWGYPYGGYGYG-YGYGYGYGDQGFSAVGYPGYYGY-----120
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 233 NGYGGGPGGNGGSPGYGGGGRGYGGGGPGYGNCGGYYDNYGGGNYGSNNDFG 292
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 ----YPGSGY----GGYGS--YGSYSGSGSYTPNVNASAG 151
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 293 NYNQPSNYGPMSKSNFGSRMGGPVGGINYPGSGSGS 333
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT      8
US-08-538-711A-7
Sequence 7, Application US/08538711A
Patent No. 5994062
GENERAL INFORMATION:
APPLICANT: MULSHINE, JAMES, L.
TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
TITLE OF INVENTION: DNA THEREOF FOR USE IN EARLY CANCER DETECTION
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/538,711A
FILING DATE: 02-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4201

```

[illegible]

Query match 17.98; Score 152.5; DB 2; Length 907;

Thu Jan 17 08:02:51 2002

us-09-554-547-16.ra1

Page 8

Search completed: January 16, 2002, 14:55:28
Job time: 320 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 14:53:38 ; Search time 14.33 Seconds
(without alignments)

818.624 Million cell updates/sec

Title: US-09-554-547-16

Perfect score: 851
Sequence: 1 MKAFFVLSTLSTALTNAR.....GSYGSYTPNVASAGAAA 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	210.5	24.7	159	2 C49773	ecdysone-dependent RNA/ssDNA-binding
2	198.5	23.3	589	2 S71954	glycine-rich prote
3	191.5	22.5	200	2 S10334	hypothetical prote
4	190.5	22.4	259	2 T15126	nucleolar protein
5	190.5	22.4	414	2 JN0866	keratin 9, type I,
6	190.5	22.4	622	2 I37984	heterogeneous ribo
7	187	22.0	373	2 S40777	heterogeneous nucl
8	183.5	21.6	345	1 B41732	glycine-rich cell
9	183	21.5	465	1 S01820	glycine-rich cell
10	180.5	21.2	165	1 KNR2G1	keratin 3, type I,
11	179	21.0	486	1 KRXL	protein RNA-bindin
12	178.5	21.0	289	2 F96770	eggshell protein p
13	178	20.9	220	2 A44805	ribonucleoprotein
14	178	20.9	385	2 S40778	nuclear DNA helica
15	176.5	20.7	1287	2 I46032	GCR 17 protein - f
16	176	20.7	196	2 E84468	probable glycine-r
17	175	20.6	154	2 E84468	keratin 1, type II
18	174	20.4	643	1 KRNU2	hypothetical prote
19	172.5	20.3	255	2 B84777	glycine-rich cell
20	172	20.2	183	1 KNR2G2	glycine-rich RNA-b
21	171.5	20.2	165	2 T03583	glycine-rich prote
22	171.5	20.2	167	2 S71779	hypothetical prote
23	170.5	20.0	252	1 S01821	RNA binding protei
24	169.5	19.9	139	2 T33968	fibroin - Chinese
25	169.5	19.9	173	2 S53050	glycine-rich prote
26	169.5	19.9	2639	2 T33328	glycine-rich prote
27	169	19.9	291	1 S31415	glycine-rich prote
28	168.5	19.8	214	1 KNRW25	keratin 10, type I
29	167	19.6	593	1 KRNU0	

30	167	19.6	633	2 H84854	probable ATP-depen
31	167	19.6	1585	2 T31611	hypothetical prote
32	166.5	19.6	346	2 S40775	ribonucleoprotein
33	166.5	19.6	404	2 B84745	probable RNA-bind
34	166	19.5	155	1 KRCHS	keratin, scale - c
35	165.5	19.4	273	2 T40618	probable cell wall
36	165.5	19.4	365	2 T04346	heterogeneous ribo
37	165	19.3	227	2 T15772	glycine-rich RNA-b
38	164	19.2	102	2 T20557	hypothetical prote
39	163	19.2	142	2 S12311	glycine-rich RNA-b
40	163	19.2	353	2 B34504	heterogeneous nucl
41	163	19.2	338	2 KNU0	glycine-rich cell
42	162.5	19.1	358	1 KNRMS2	keratin, type II c
43	162.5	19.1	581	1 KRMS2	glycine-rich prote
44	161.5	19.0	203	1 J01061	glycine-rich prote
45	161	18.9	169	1 S38331	glycine-rich RNA-b

ALIGNMENTS

```
RESULT 1
C49773
ecdysone-dependent cytochrome-like protein EDG-91 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #ext_change 24-Nov-1999
C:Accession: C49773
R:Apple, R.T., Pristrom, J.W.
Dev. Biol. 146, 569-582, 1991
A:Title: 20-Hydroxyecdysone is required for, and negatively regulates, transcription
A:Reference number: A49773; MUID:91323677
A:Accession: C49773
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <APP>
A:Cross-references: GB:M71250; NID:9157326; PID:9157327; GB:M71251; NID:9157328; PID:
C:Genetics:
A:Gene: FlyBase:Edg91
A:Cross-references: FlyBase:FBgn0004554
C:Superfamily: glycine-rich cell wall structural protein 3

Query Match          24.7%; Score 210.5; DB 2; Length 159;
Best Local Similarity 42.8%; Pred. No. 1.4e-09;
Matches 59; Conservative 6; Mismatches 34; Indels 39; Gaps 10;

QY 38 LVAGIERAGPRGYPLTAS-IGSEVGARLGGRAAGVSS-----YGYGP---SMGYPY 87
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 14 LINGGQOAPVKTGRTLLGLGGFGSGVGLSAGIGVGGLYSGFGGCGVPGGYASGYP- 72

QY 88 GGY-GGYGGYGGYGYDQGFSAV-----GGY-----PGYGYYPYPSG-----YGGYGG 132
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 73 GGYGGYSGYNGYGG--SGFGGYYPGGYSGFGHPRHGYGGYYPGGYSGYHNOGGSYGGH 130

QY 133 Y-----GSGYGG 139
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 131 YSOSQYSNGYGGGYGG 148

RESULT 2
S71954
RNA/ssDNA-binding protein TAFII68 - human
N:Alternate names: TAFII-binding protein-associated factor TAFII68
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #ext_change 20-Jun-2000
C:Accession: S71954
R:Bercolotti, A.; Lutz, Y.; Heard, D.J.; Chambon, P.; Tora, L.
EMBO J. 15, 5022-5031, 1996
A:Title: hTAFII68, a novel RNA/ssDNA-binding protein with homology to the pro-oncopro
A:Reference number: S71954; MUID:97045110
A:Accession: S71954
A:Molecule type: mRNA
A:Residues: 1-589 <BER>
```

A:Cross-references: EMBL:X88893; NID:g1628402; PIDN:CAA67398.1; PID:g1628403
 A:Experimental source: HeLa cells
 C:Genetics:
 A:Gene: taf68
 C:Complex: up to 13 TAFn-binding protein-associated factors (TAFII)s form together with
 C:Function:
 A:Description: required for activated transcription; binds RNA and single stranded DNA;
 C:Superfamily: RNA-binding protein EMS; ribonucleoprotein repeat homology
 C:Keywords: RNA binding
 F:232-307/Domain: ribonucleoprotein repeat homology <RPM>

Query Match 23.3%; Score 198.5; DB 2; Length 589;
 Best Local Similarity 43.5%; Pred. No. 3; 7e-08;
 Matches 60; Conservative 3; Mismatches 52; Indels 23; Gaps 9;

QY 20 RAGRLGSDLDTFGRVHNLVAGIERAGPRGPGTLASTIGEVGARLGG-RAGVGVSSYGY 78
 Db 456 RGGGCGD-----RGGGCGG-DRGG-GYGDRGGYGGGGRGGGDRGGYGGDRGGY 505
 QY 79 GYPSMGY--PYGYYGGYGGYGGYDGFSGAYG--YPGYGYVYPPSGY---GGGYG 130
 Db 506 GGDGCGYGGDRGGYGGDRSGRNGYGG-DRGGSSYGGDRSGYGGDRSGGTYGGDRGGY 564
 QY 131 ---GSYGGSYGGSYTYPN 145
 Db 565 GDRGGYGGKMGGRNDYRN 582

RESULT 3
 S10334
 glycine-rich protein precursor - barley
 C:Species: Hordeum vulgare (barley)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S10334
 R:Rehde, W.; Rosch, K.; Kroege, K.; Salami, F.
 Plant Mol. Biol. 14, 1057-1059, 1990.
 A:Title: Nucleotide sequence of a Hordeum vulgare gene encoding a glycine-rich protein
 A:Reference number: S10334; MIMD:91346692
 A:Accession: S10334
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-200 <RHS>
 A:Cross-references: EMBL:X52580; NID:g18995; PIDN:CAA36811.1; PID:g295808
 C:Genetics:
 A:Introns: 29/1
 C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 22.5%; Score 191.5; DB 2; Length 200;
 Best Local Similarity 36.5%; Pred. No. 5e-08;
 Matches 65; Conservative 9; Mismatches 59; Indels 45; Gaps 9;
 QY 2 KAFVLSLSTALITNA-----ARAGRLGSDLDTF---GRVHNLVAGIERAGRY 50
 Db 6 KGLVVALLLAAAILVSADEHPQAKKEENFHHGGGHHGKRG---HGGGY 62
 QY 51 PGLTASIGEVGARLGGYAGVSSYGYPPSMGYP-----YGYGGY----- 93
 Db 63 GG-----GGYGGGGGYPGGG---GGYGGGGGYPCHGEGGGGGYGGGGYPCHEGEG 114
 QY 94 GYGYGGYDGFSGAYGYPGYVYPPSGYGGYGGYGGYGGSYTYPNVARSAG 151
 Db 115 GYGGGGGHHGHEGGGGYGGGGYH---CHGEGGGGGYGGGGG---YPGHGGGG 166

RESULT 4
 T15126
 hypothetical protein T20B6.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
 C:Accession: T15126
 R:Beck, C.; Wamsley, P.

submitted to the EMBL Data Library, April 1997
 A:Description: The sequence of C. elegans cosmid T20B6.
 A:Reference number: Z18297
 A:Accession: T15126
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-259 <BEC>
 A:Cross-references: EMBL:AF000193; NID:g1946986; PID:g1946988; PIDN:AMB52890.1; GSPDB
 A:Experimental source: strain Bristol N2; clone T20B6
 C:Genetics:
 A:Gene: CESP:T20B6.3
 A:Map position: 3/2
 A:Introns: 9/2; 231/1
 C:Superfamily: Phaseolus glycine-rich protein 1.0

Query Match 22.4%; Score 190.5; DB 2; Length 259;
 Best Local Similarity 42.2%; Pred. No. 7; 4e-08;
 Matches 57; Conservative 7; Mismatches 46; Indels 25; Gaps 9;

QY 13 AALFNA-ARAGRLGSDLDTFGRVHNLVAGIERAGPRGPGTLASTIGEVGARLGGY 71
 Db 68 AAMTSTGRGGYGGSGGATGFGNGY-----GGGCGG-----GGDRGGMG-CGY 114
 QY 72 GVSSYGYGYPGSMGYPYGYYGGYGGYGGYDGFSGAYGYPGYVYPPSGY----- 125
 Db 115 GGGYGGGGDGGG-GYGGYG-GYGGMGGGPGGYGMYGGGGGGGDF--GGYGGG 170
 QY 126 -GGYGGSYGSGYCG 139
 Db 171 GGGYGGGGDGGYGG 185

RESULT 5
 JN0866
 nucleolar protein NOP3 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: mitochondrial targeting suppressor 1; NUP3 protein; protein YDR432
 C:Species: Saccharomyces cerevisiae
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 23-Mar-2001
 C:Accession: JN0866; A44401; S42161; S69713; S27435; S31322
 R:Ellis, E.M.; Reid, G.A.
 Gene 132, 175-183, 1993
 A:Title: The Saccharomyces cerevisiae NUP3 gene encodes a putative RNA-binding protein
 A:Reference number: JN0866; MIMD:94040759
 A:Accession: JN0866
 A:Molecule type: DNA
 A:Residues: 1-414 <EHL>
 A:Cross-references: EMBL:X70951; NID:g288585; PIDN:CAA50291.1; PID:g288587
 R:Russell, I.D.; Tollervey, D.
 J. Cell Biol. 119, 737-747, 1992
 A:Title: NOP3 is an essential yeast protein which is required for pre-rRNA processing
 A:Reference number: A44401; MIMD:93054908
 A:Accession: A44401
 A:Molecule type: DNA
 A:Residues: 1-414 <RUS>
 A:Cross-references: EMBL:X66019; NID:g4039; PIDN:CAA66817.1; PID:g4040
 A:Note: sequence extracted from NCBI backbone (NCBP:117151)
 R:Bosste, M.A.; Dehoratius, C.; Barcelo, G.; Silver, P.
 MOL. BIOL. Cell 3, 875-893, 1992
 A:Title: A mutant nuclear protein with similarity to RNA binding proteins interferes
 A:Reference number: S42161; MIMD:93004942
 A:Accession: S42161
 A:Molecule type: DNA
 A:Residues: 1-414 <ROS>
 A:Cross-references: EMBL:M66731; NID:g172051; PIDN:AAA34818.1; PID:g172052
 A:Note: the authors translated the codon GAA for residue 121 as Gln and GAA for resid
 R:Diehl, F.S.
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and
 A:Reference number: S69555
 A:Accession: S69713
 A:Molecule type: DNA
 A:Residues: 1-414 <DIE>

```

A:Cross-references: EMBL:U03007; NID:g927685; PIDN:AA64865.1; PID:g927659; MIPS:YDR432w
C:Genetics:
A:Gene: SGD:NPL3; NOP3; MTS1
A:Cross-references: SGD:S0002840; MIPS:YDR432w
A:Map position: 4R
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C:Keywords: nucleolus; nucleus; RNA binding
F:126-185/Domain: ribonucleoprotein repeat homology <RNM1>
F:201-265/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match      22.4%   Score 190.5; DB 2; Length 414;
Best Local Similarity 45.18; Pred. No. 1.1e-07;
Matches 55; Conservative 4; Mismatches 46; Indels 17; Gaps 7;

OY  42  IERAGPGYGLTASIGCEGARLGG--RAGGVSSYGYPSMGYPYGY--GGYGY- 96
      | : | : | | | | | | | | | | | | | | | | | | | | | | |
DB  279  IRRSNMGCFRGRGFRGCGFRGCGFRGCGFRGCGFRGCGFRGCGFRGCG- 336
      | : | : | | | | | | | | | | | | | | | | | | | | | | |
OY  97  -GGYGYDGGFGSAGYGYPGYGYVPSGYG--GYG--GSYGSYGSYTYPNVRAS 149
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  337  RGGYGSRRGYDPSRGGYDPSRGGYSGRGGYSGPRNDYPPRGSYGSRCG---YDPRGD 393
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  150  AG 151
      |
DB  394  YG 395

RESULT 6
137984
Keratin 9, type I, cytoskeletal - human
N:Alternate names: cytokeratin 9; scatter protein 60K chain, placental
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence,revision 02-Jul-1996 #text,change 29-Oct-1999
C:Accession: I37984; S40307; S77921; S41161; B35494; I37943
R:Reis, A.; Hemmes, H.C.; Langbein, L.; Digeed, M.; Mischke, D.; Drechsler, M.; Schroed
N:Gene: 6, 174-179, 1994
A:Title: Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK).
A:Reference number: I37984; MUID:94214498
A:Accession: I37984
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-622 <RES>
A:Cross-references: EMBL:X75015; NID:9453154; PIDN:CAA52924.1; PID:9453155
R:Langbein, L.; Held, H.W.; Moll, I.; Franke, W.W.
Differenciation 55, 57-71, 1993
A:Title: Molecular characterization of the body site-specific human epidermal cytokerati
A:Reference number: I37943; MUID:94131202
A:Accession: S40307
A:Molecule type: mRNA
A:Residues: 1-622 <LNA>
A:Cross-references: EMBL:Z29074; NID:9435475; PIDN:CAA82315.1; PID:9435476
A:Accession: S77921
A:Molecule type: protein
R:Langbein, L.
submitted to the EMBL Data Library, December 1993
A:Reference number: S41161
A:Accession: S41161
A:Molecule type: mRNA
A:Residues: 1-11, 'SR', 13-622 <LNA>
A:Cross-references: EMBL:Z29074; NID:9435475; PIDN:CAA82315.1; PID:9435476
Biochem. Biophys. Res. Commun. 168, 1082-1088, 1990
A:Title: Human placenta contains an epithelial scatter protein.
A:Reference number: A35494; MUID:90267446
A:Accession: B35494
A:Molecule type: protein
A:Residues: 'X', 450-465 <ROS>
C:Genetics:
A:Gene: GDB:KRT9; EPPK
A:Cross-references: GDB:303970; OMIM:144200
A:Map position: 17q12-17q21

```

A: Introns: 213/3; 241/2; 293/3; 347/3; 389/3; 464/2
A>Note: defects in this gene may cause epidermolytic palmoplantar keratoderma
C: Superfamily: cytoskeletal keratin
C: Keywords: coiled coil; intermediate filament
F: 1-153/Domain: head #status predicted <HNA>
F: 154-459/Domain: helical rod #status predicted <ROD>
F: 460-622/Domain: tail #status predicted <TAI>

Query Match 22.4%; Score 190.5; DB 2; Length 622;
Best Local Similarity 37.4%; Pred. No. 1.6e-07;
Matches 52; Conservative 18; Mismatches 42; Indels 27; Gaps 7

Oy 11 SPVALTNAARAGRLGSLDTFGRRVHGNIYAGIERAPRGYPGLITASIGSEVG--ARLGR 68
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 7 SSSYLTS---GGGGGGGLSGSGSIRSS-YSRSSSCGRGCGRFRSSSSGYGGSSRVCR 62
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Oy 69 AGVGVSVSYGTGPMSKVPRTGYGGYGXY-----GGYGTDDGFSAHYGTYGYXYP 122
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 63 GGCG-----SPFYSYG-G-SGGGFSASSLGGEGGGSGRCFGASG-----GYSS 107
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Oy 123 SGYGGGYGSGSYGSGY 141
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 108 GGFGGGFGGSGGFGG 126

RESULT 7
SA0777

heterogeneous ribonucleic particle protein A3 - African clawed frog
N: Alternate names: heterogeneous nuclear ribonucleoprotein A3
C: Species: Xenopus laevis (African clawed frog)
C: Date: 31-Dec-1993 #sequence__revision 02-Aug-1994 #text_change 23-Jul-1999
C: Accession: SA40777; SA1738
R: Good, P.J.; Rebberdt, M.L.; David, I.B.
Nucleic Acids Res. 21, 999-1006, 1993
A: Title: Three new members of the RNP protein family in Xenopus.
A: Reference number: SA40774; MUID: 93197168
A: Accession: SA40777
A: Status: nucleic acid sequence not shown
A: Molecule type: mRNA
A: Residues: 1-373 <GOOL>
A: Cross-references: EMBL: L02956
R: Good, P.J.; Lai, M.; Rebberdt, M.L.; David, I.B.
submitted to the EMBL Data Library, October 1992
A: Reference number: SA1738
A: Accession: SA1738
A: Molecule type: mRNA
A: Residues: 1-85, 'S', '87-346, 'S', '348-373 <GOOZ>
A: Cross-references: EMBL: L02956; NID: g214744; PIDN: AAA49949.1; PID: g214745
C: Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
F: 28-94/Domain: ribonucleoprotein repeat homology <RM1>
F: 119-185/Domain: ribonucleoprotein Repeat homology <RM2>

Query Match 22.0%; Score 187; DB 2; Length 373;
Best Local Similarity 32.8%; Pred. No. 1.9e-07;
Matches 57; Conservative 8; Mismatches 55; Indels 54; Gaps 6

Oy 16 TMAARAGRLGSLDTFGRRVHGNIYAGIERAG-PRTGPGLTASTIGEVARILGGRAGVYS 74
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 201 TSAQRGGGGSNFMGR--GGNYGGGDGNFGRGGGGFGNRNGYGC--GARGGGGYG 256
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Oy 75 SYGYGPSMGTYRYGCG-----CYGTYGTYGTYDGFSAHYGTYPRY----- 116
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 257 GGGDGNGGCGDGNYGPGPYGGRYGGSPGYGNQGGYGGGGGYYDYDNESGNFGGN 316
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Oy 117 ---YGY-----YPSGYGGGSGSYGSGY 139
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 317 YNDPFGYGGQQSNYCPMKGSGSFSGRSSGSGRSGRPYGGGTSGSGGGGGGSGTG 370

RESULT 8
BA1732

heterogeneous nuclear RNP protein clone pHRP40.2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: B41732
R:Matunis, E.L.; Matunis, M.J.; Dreyfuss, G.
J. Cell Biol. 116: 257-269, 1992
A:Title: Characterization of the major hnRNP proteins from Drosophila melanogaster.
A:Reference number: A41732; MUID:92112968
A:Accession: B41732
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-345 <MAT>
A:Cross-references: GB:X62638; NID:q11039; PIDN:CAAA4504.1; PID:q11040
A>Note: Sequence extracted from NCBI Backbone (NCBIN:76632, NCBIPI:76633)
C:Genetics:
A:Gene: FlyBase:sgd
A:Cross-references: FlyBase:FBgn0003498
A:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
F:57-123/Domain: ribonucleoprotein repeat homology <RRM1>
F:137-203/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 21.6%; Score 183.5; DB 1; Length 345;
Best Local Similarity 45.6%; Pred. No. 3.2e-07;
Matches 52; Conservative 5; Mismatches 30; Indels 27; Gaps 9;

QY 37 NLVAGIERAGPRGYPGTLASIGEGVAGRLGAGV----GVSSY-GYGYPSWGYPGY 90
Db 217 NQMMGMGMGPRG--GMNGGRG-----YGRGGYNNQMDQSGSGG-----GCG 261
QY 91 GGY--GGYGGT--GGYDGFSGANGYGYGYPPSGY--GGYGGSGSGSGY 139
Db 262 GGYAGGYGDIYAGGYNGYDYGDY--GYGGGEGMNGYGGGGGGMNGGGRG 314

RESULT 9
S01820
glycine-rich cell wall protein 1.8 precursor - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 30-Sep-1989 #sequence_revision 19-May-1994 #text_change 16-Jul-1999
C:Accession: S01820
R:Keller, B.; Sauer, N.; Lamb, C.J.
EMBO J. 7, 3625-3633, 1988
A:Title: Glycine-rich cell wall proteins in bean: gene structure and association of the
A:Reference number: S01820; MUID:89091109
A:Accession: S01820
A:Molecule type: DNA
A:Residues: 1-465 <KEL>
A:Cross-references: EMBL:X13596; NID:921002; PIDN:CAA31932.1; PID:q21003
C:Comment: This protein is enriched in the cell wall fraction of young hypocotyls and ov
lys.
C:Comment: Much of the sequence consists of tandemly repeated 22-residue segments with
C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8
C:Keywords: cell wall; structural protein; tandem repeat
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-465/Product: glycine-rich cell wall protein 1.8 #status predicted <MAT>

Query Match 21.5%; Score 183; DB 1; Length 465;
Best Local Similarity 39.4%; Pred. No. 4.6e-07;
Matches 52; Conservative 7; Mismatches 59; Indels 14; Gaps 4;

QY 20 RAGRLGSDLDFGRHGNLVAGIERAGPRGYPGTLASIGEGVAGRLGAGVSSYGY 79
Db 323 QGGGAGGGAAYVGE-HGGGYGGGGGGGGGGGGGGGGGGGGGGGAGG---YGTG 378
QY 80 YPSNGYPGYGGYGGYGGY-----GYGYDGFSGANGY--PGYGYPPSGYGGY 129
Db 379 GEHGGYGG 438
QY 130 GGSYGGSGYGSY 141
Db 439 GYGAGGAHGGY 450

RESULT 10
KNR261
glycine-rich cell wall structural protein 1 precursor (clone lambda-313) - rice
C:Species: Oryza sativa (rice)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C:Accession: S13385
R:Lei, M.; Wu, R.
Plant Mol. Biol. 16, 187-198, 1991
A:Title: A novel glycine-rich cell wall protein gene in rice.
A:Reference number: S13385; MUID:91370862
A:Accession: S13385
A:Molecule type: DNA
A:Residues: 1-165 <LEI>
A:Cross-references: EMBL:X53596; NID:920246; PIDN:CAA37665.1; PID:g20247
C:Genetics:
A:Gene: grp-1
C:Superfamily: glycine-rich cell wall structural protein 1
C:Keywords: cell wall; duplication; structural protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-165/Product: glycine-rich cell wall structural protein 1 #status predicted <MAT>
F:30-55/Region: repeat R1
F:56-62/Region: repeat R2
F:62-92/Region: repeat R1
F:93-99/Region: repeat R2
F:100-131/Region: repeat R1
F:132-138/Region: repeat R2
F:139-160/Region: repeat R1

Query Match 21.2%; Score 180.5; DB 1; Length 165;
Best Local Similarity 37.5%; Pred. No. 2.9e-07;
Matches 60; Conservative 10; Mismatches 55; Indels 35; Gaps 6;

QY 5 FVLSLSTALITNAR-----ACRLGSDLDFGRHGNLV--AGIERA 45
Db 10 FLILLTSLSKSNARVIRKNGSGGGGGGGGGGGGGSGSGGYGKRGCGSG 69
QY 46 GPRGYPGTLASIGEGVAGRLGAGVSSYGY-----PSWGYPGYGGYGGYGGY 100
Db 70 GGQSGG-----GGGGGGGNGSGSG--SGYGYGGGNGGAGCGGGGGGGGGGG 123
QY 101 GYDGFSGANGYGYGYPPSGYGGGGGGGGGGGGSGSGSGS 140
Db 124 GSGGSGSGY-----YCKRGGGGGGGGGGGGGGGGGGG 158

RESULT 11
KRXL
keratin 3, type I, cytoskeletal (clone pUF1001) - African clawed frog
N:Alternate names: cytokeratin, 51K
C:Species: Xenopus laevis (African clawed frog)
C>Date: 28-Aug-1985 #sequence_revision 30-Sep-1991 #text_change 10-Dec-1999
C:Accession: S01327; A02943
R:Hoffmann, W.; Sterrer, S.; Koentgen, S.; Koentgen, S.
FEBS Lett. 237, 178-182, 1988
A:Title: Amino acid sequence microheterogeneities of a type I cytokeratin of M(r) 510
A:Reference number: S01327; MUID:89005622
A:Accession: S01327
A:Molecule type: mRNA
A:Residues: 1-486 <HO>
A:Cross-references: EMBL:Y00968; NID:964490; PIDN:CAA68783.1; PID:g64491
R:Hoffmann, W.; Praet, J.K.
EMBO J. 3, 1301-1306, 1984
A:Title: Amino acid sequence of the carboxy-terminal part of an acidic type I cytoke
A:Reference number: A02943; MUID:84261417
A:Accession: A02943
A:Molecule type: mRNA
A:Residues: 339-351, 'T', 353-486 <HO>
A:Cross-references: GB:X00629; NID:964868; PIDN:CAA25263.1; PID:g833626
A>Note: this sequence has been revised in paper S01327
C:Superfamily: cytoskeletal keratin

I46032
nuclear DNA helicase II - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 02-Feb-2001
C:Accession: I46032; S49822
R:Zhang, S.; Maacke, H.; Grosse, F.
J. Biol. Chem. 270, 16422-16427, 1995
A:Title: Molecular cloning of the gene encoding nuclear DNA helicase II. A bovine homolo
A:Reference number: I46032; MUID:95332357
A:Accession: I46032
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1287 <2HA>
A:Cross-references: EMBL:X82829; NID:9577738; PIDN:CAA58036.L; PID:9577739
C:Genetics:
A:Gene: ndh2
C:Keywords: ATP; nucleotide binding; P-loop
F:408-415/Region: nucleotide-binding motif A (P-loop)
F:504-509/Region: nucleotide-binding motif B
F:508-511/Region: DEXH motif

Query Match 20.78; Score 176.5; DB 2; Length 1287;
Best Local Similarity 37.9%; Pred. No. 3.5e-06;
Matches 53; Conservative 1; Mismatches 45; Indels 41; Gaps 5;
OY 37 NLYAGIERAGPRGYPGULTASIGFEVAGRLGGRAGVSSY-GGYPSMGYPYGGYGGY 95
DB 1144 NLMGTRRYGDPKPPKMARVDSGYRRGG-----SSYSGGGLGGLGTYGTGGYGGGG 1197
OY 96 YGYGYGYDOG-----FGSAY-----GGYPGYGYGYPPSGYGGG 128
DB 1198 YGGRGYSGGGGSGNSFRSGYVGGGGYGGGGGFRGLSRGGRGMGSDYRGESGCG 1257
OY 129 YGGSYG-----GSYGSY 141
DB 1258 YRSGSGFQRGGRGCGYGGGY 1277

Search completed: January 16, 2002, 14:55:49
Job time: 131 sec

Db 14 LIAGGGAAPYKTEGRRLGLLGGFGSGVLSAGIGVGGGLYSGFGGPGGYASGXP 72
 Qy 88 GGY-GGYGGYGGYDQGFSAV---GGY-----PGYGYPPSG-----YGGYGGG 132
 Db 73 GYGCGGYSYNGYGG---SGFGGYPPGGGYSFGFHRPHYGHGYPGGGSHNMGGSYGGH 130
 Qy 133 Y-----GGSYGG 139
 Db 131 YSQSQYNSNGYGGGCGG 148

RESULT 2
 RB56_HUMAN STANDARD; PRT; 592 AA.
 ID Q92804; Q92751;
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TATA-BINDING PROTEIN ASSOCIATED FACTOR 2N (RNA-BINDING PROTEIN 56)
 DE TAFII168 (TAFII168).
 GN TAF2N OR RB56.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=97124846; PubMed=8954779;
 RA Morohoshi F., Arai K., Takahashi E.-I., Tanigami A., Ohki M.;
 RT "Cloning and mapping of a human RBp56 gene encoding a putative RNA
 RT binding protein similar to FUS/TLS and EWS proteins.";
 RL Genomics 38:51-57(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM), AND SEQUENCE OF 282-297 & 307-320.
 RX MEDLINE=97045110; PubMed=8890175;
 RA Bertolotti A., Lutz Y., Heard D.J., Chambon P., Tora L.;
 RT "TAFII168, a novel RNA/ssDNA-binding protein with homology to the
 RT pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA
 RT polymerase II.";
 RL EMO J. 15:5022-5031(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=99013873; PubMed=9795213;
 RA Morohoshi F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,
 RA Munakata M., Ohki M.;
 RT "Genomic structure of the human RBp56/TAFII68 and FUS/TLS genes.";
 RL Gene 221:191-198(1998).
 CC -|- FUNCTION: RNA AND ssDNA-BINDING PROTEIN THAT MAY PLAY SPECIFIC
 CC ROLES DURING TRANSCRIPTION INITIATION AT DISTINCT PROMOTERS. CAN
 CC ENTER THE PREINITIATION COMPLEX TOGETHER WITH THE RNA POLYMERASE
 CC II (POL. II).
 CC -|- SUBUNIT: BELONGS TO THE RNA POLYMERASE II (POL. II) TRANSCRIPTOMAL
 CC MULTIPROTEIN COMPLEX, TOGETHER WITH THE TATA-BINDING PROTEIN (TBP)
 CC AND OTHER TBP-ASSOCIATED FACTORS (TAFIIIS).
 CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM (SHOWN HERE)
 CC AND A SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -|- TISSUE SPECIFICITY: UBIQUITOUS. OBSERVED IN ALL FETAL AND ADULT
 CC TISSUES.
 CC -|- DISEASE: A FORM OF EXTRASKELETAL MIXOID CHONDROSARCOMAS (EMC) IS
 CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(9;17)(Q22;Q11) THAT
 CC INVOLVES TAF2N AND NR4A3.
 CC -|- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -|- SIMILARITY: CONTAINS 1 RANBP-TYPE ZINC FINGER.
 CC -|- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; U51334; AAC50932.1; -;
 DR EMBL; X98893; CA67398.1; -;
 DR EMBL; AB010067; BAA33811.1; -;
 DR EMBL; AB010057; BAA33811.1; JOINED.
 DR EMBL; AB010058; BAA33811.1; JOINED.
 DR EMBL; AB010059; BAA33811.1; JOINED.
 DR EMBL; AB010060; BAA33811.1; JOINED.
 DR EMBL; AB010061; BAA33811.1; JOINED.
 DR EMBL; AB010062; BAA33811.1; JOINED.
 DR EMBL; AB010063; BAA33811.1; JOINED.
 DR EMBL; AB010064; BAA33811.1; JOINED.
 DR EMBL; AB010065; BAA33811.1; JOINED.
 DR EMBL; AB010066; BAA33811.1; JOINED.
 DR EMBL; AB010067; BAA33812.1; -;
 DR EMBL; AB010057; BAA33812.1; JOINED.
 DR EMBL; AB010058; BAA33812.1; JOINED.
 DR EMBL; AB010059; BAA33812.1; JOINED.
 DR EMBL; AB010060; BAA33812.1; JOINED.
 DR EMBL; AB010061; BAA33812.1; JOINED.
 DR EMBL; AB010062; BAA33812.1; JOINED.
 DR EMBL; AB010063; BAA33812.1; JOINED.
 DR EMBL; AB010064; BAA33812.1; JOINED.
 DR EMBL; AB010065; BAA33812.1; JOINED.
 DR EMBL; AB010066; BAA33812.1; JOINED.
 DR MIM; 601574; -;
 DR InterPro; IPR000504; RRM.
 DR InterPro; IPR001876; Znf-RanBP.
 DR Pfam; PRF0076; rrm; 1.
 DR Pfam; PRF0641; zf-RanBP; 1.
 DR SMART; SM00360; RRM; 1.
 DR SMART; SM00547; Znf_RB2; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 KW Proto-oncogene: RNA-binding; DNA-binding; Nuclear protein; Repeat;
 KW Alternative splicing; Chromosomal translocation; Zinc finger; Zinc;
 KW Metal-binding.
 FT DOMAIN 1 208 GLN/GLY/SER/TYR-RICH.
 FT 234 320 RNA-BINDING (RRM).
 FT 320 590 ARG/GLY-RICH.
 FT ZN_FING 360 379 C4-TYPE (POTENTIAL).
 FT DOMAIN 407 575 21 X APPROXIMATE TANDEM REPEATS OF D-R-
 FT {S,G}(0,3)-G-G-Y-G-G.
 FT REPEAT 407 413 1.
 FT REPEAT 414 420 2.
 FT REPEAT 421 429 3.
 FT REPEAT 430 439 4.
 FT REPEAT 440 448 5.
 FT REPEAT 449 457 6.
 FT REPEAT 458 465 7.
 FT REPEAT 466 473 8.
 FT REPEAT 474 481 9.
 FT REPEAT 482 488 10.
 FT REPEAT 489 496 11.
 FT REPEAT 497 503 12.
 FT REPEAT 504 510 13.
 FT REPEAT 511 517 14.
 FT REPEAT 518 524 15.
 FT REPEAT 525 533 16.
 FT REPEAT 534 543 17.
 FT REPEAT 544 551 18.
 FT REPEAT 552 560 19.
 FT REPEAT 561 568 20.
 FT REPEAT 569 575 21.
 FT VARSPLIC 60 62 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 592 AA; 61830 MW; 73D37C171E1E2BCA CRC64;

Qy 20 RAGRLGSDLDTPFGRVHGNLYAGIERAGPRGYPGLTASIGGEVAGRLGG-RAGGVSSYGY 78
 Query Match 23.3%; Score 198.5; DB 1; Length 592;
 Best Local Similarity 43.5%; Pred. No. 2.6e-07;
 Matches 60; Conservative 3; Mismatches 52; Indels 23; Gaps 9;


```

RT      "A novel glycine-rich cell wall protein gene inrice.";
CC      Plant Mol. Biol. 16:187-198(1991).
CC      -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
CC      -1- SUBCELLULAR LOCATION: CELL WALL (POTENTIAL).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X53596; CAA37665.1; -.
DR      PIR: S13385; KNR2G1.
DR      HSSP: P30129; 4DPV.
CC      Cell wall; Structural protein; Repeat; Signal.
FT      SIGNAL          1      23      POTENTIAL.
FT      KW              1      23      GLYCINE-RICH CELL WALL STRUCTURAL
FT      CHAIN           24      165      PROTEIN 1.
FT      FT              31      159      GLY-RICH.
FT      REPEAT          56      62      R2 (TYR-RICH).
FT      REPEAT          93      99      R2 (TYR-RICH).
FT      REPEAT          132     138      R2 (TYR-RICH).
SO      SEQUENCE       165 AA; 13536 MW; E36CE31C3650AC9A CNC64;

Query Match          21.2%; Score 180.5; DB 1; Length 165;
Best Local Similarity 37.5%; Pred. No. 1.6e-06;
Matches 60; Conservative 10; Mismatches 55; Indels 35; Gaps 6;

QY      5 FVLSLSTALTNAR-----ACRLSDLDTFERVHGNLY--AGIERA 45
DB      10 FLTLTLTSLKSNMARVRYKNGSGGGGGGGGGGGSGSGGYGKAGGSG 69
QY      46 GPRGPGGLTASIGCGVGARLGRAGVGVSSVYGYG----PSWGYRPGYGGYGGYGGV 100
DB      70 GGGGSGG-----GGGGGGGGGGNGSSSG--SGYGYGGGNGGAGGGGGGGGGGGGGG 123
QY      101 GYDQFGSAYGYPGYGYGYPYPGYPGGYGGYGGSGSYGSS 140
DB      124 GSGGGSGSGYG-----YGYKGGGGGGGGGGGGGGGGG 158

RESULT 10
DDX9_MOUSE STANDARD: PRT: 1380 AA.
ID      DDX9_MOUSE STRAND:
AC      070133; 035931; 054703;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      ATP-DEPENDENT RNA HELICASE A (NUCLEAR DNA HELICASE II) (NDH II)
DE      (DEAD-BOX PROTEIN 9) (WHEL-5).
GN      DDX9.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=128/SV;
RC      MEDLINE=98149984; PubMed=9480750;
RA      Lee C.-G., Ekl T., Okumura K., da Costa Soares V., Hurwitz J.;
RT      "Molecular analysis of the cdna and genomic DNA encoding mouse RNA
RL      helicase A.";
RL      Genomics 47:365-371(1998).
RP      [2]
RP      SEQUENCE OF 386-919 FROM N.A.
RC      STRAIN=C57BL/6;
RC      Kistelow P., Miazek A.;
RA      Submitted (Mar-1997) to the EMBL/GenBank/DBJ databases.
RT      -1- FUNCTION: UNKNOW DOUBLE-STRANDED DNA AND RNA IN A 3' TO 5'
CC      DIRECTION. GENERATES MULTIPLE SECONDARY STRUCTURES THAT INFLUENCE

```

CC	RAA-BINDING PROTEINS.	MAY PLAY A ROLE IN X-LINKED GENE EXPRESSION.
CC	- - SUBCELLULAR LOCATION:	NUCLEAR.
CC	- - SIMILARITY:	BELONGS TO THE "DEAD" BOX FAMILY HELICASE. DEAH SUBFAMILY.
CC	- - SIMILARITY:	CONTAINS 2 DBRM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC	This SWISS-PROT entry is copyright.	It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	EMBL; U91922;	AAC05725.1; .
DR	EMBL; AF023530;	AAC05301.1; .
DR	EMBL; U92080;	AAB72087.1; .
DR	MGI; 108177;	Ddx9.
DR	InterPro: IPR001410;	DEAD.
DR	InterPro: IPR002464;	DEAH_ATP_helicase.
DR	InterPro: IPR001159;	DS_RBD.
DR	InterPro: IPR001650;	Helicase_C.
DR	Pfam; PF00035;	dsmr; 2.
DR	Pfam; PF00271;	helicase_C; 1.
DR	SMART; SMO0487;	DEXDC; 1.
DR	SMART; SMO0358;	DSRM; 2.
DR	SMART; SMO0490;	HELICC; 1.
DR	PROSITE; PS00690;	DEAH_ATP_HELICASE; 1.
DR	PROSITE; PSS0137;	DS_RBD; 2.
KW	Helicase; RNA-binding;	DNA-binding; Repeat; Nuclear protein;
KW	ATP-binding.	
FT	DOMAIN	3 71 DIRM 1.
FT	DOMAIN	182 254 DRBM 2.
FT	NP_BIND	413 420 ATP (POTENTIAL).
FT	SITE	513 516 DEAH BOX.
FT	DOMAIN	588 597 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	DOMAIN	1171 1380 ARG/GLY/SER/TYR-RICH.
FT	CONFLICT	46 46 R -> A (IN REF. 1); AAC05301).
FT	CONFLICT	674 674 N -> M (IN REF. 2).
FT	CONFLICT	748 748 T -> I (IN REF. 2).
FT	CONFLICT	831 831 I -> V (IN REF. 2).
SQ	SEQUENCE	1380 AA; 149582 MW; 0P23755C291D2FE1 CRC64;
<hr/>		
Query Match 21.2%; Score 180; DB 1; Length 1380;		
Best Local Similarity 41.8%; Pred. No. 9.9e-06;		
Matches 59; Conservative 6; Mismatches 52; Indels 24; Gaps 9;		
OY	22 GRLSDDTFGRVHGNLYAGIERACPRIYPGILTASIGGEVGARLCGRGNGVS-SYGYCY 80	
Db	1216 GGFSGGGGGFGSGGGG-FGG-----GGGGFSG--GGGGCGFGGRGGGGGGGFGGNG 1268	
OY	81 PSMGVPGVGYGGYGGYGVDGFGSAVGGPYGVY-----YPSGGYGGYGGSYGG 135	
Db	1269 GGYGGGGGYYG-GGGGGYGGSGLYGG--GGYGGEGEYSISPNSYKRGNYGGGGGGRNG 1325	
OY	136 SYGSGYTYPN-----VRASAG 151	
Db	1326 SQGG---YRNFMGCDYRGSSG 1343	
<hr/>		
RESULT 11		
ID KICO_XENLA	STANDARD;	PRT; 486 AA.
AC P02537;		
DT 21-JUL-1986 (Rel. 01, Created)		
DT 01-NOV-1988 (Rel. 09, Last sequence update)		
DT 20-AUG-2001 (Rel. 40, Last annotation update)		
DE KERATIN 3, TYPE I CYTOSKELETAL 51 KDA (51 KDA CYTOKERATIN).		
OS Xenopus laevis (African clawed frog)		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Amphibia; Batrachia; Anura; Mesopodactylia; Pipridae; Pipidae;		
OC Xenopodinae; Xenopus.		

RL J. Biol. Chem. 269:31457-31465(1994).
 CC -1- FUNCTION: MAY REGULATE MRNA TRANSLATION AND STABILITY. IT BINDS
 CC TO POLY(A) AND POLY(U) REGIONS OF RNA. THIS BINDING IS INHIBITED
 CC WHEN THE PROTEIN IS PHOSPHORYLATED.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR. AT THE ONSET OF THE
 CC CRYPTOBlastic PHASE, WHEN TRANSCRIPTION IS BLOCKED, IT ACCUMULATES
 CC IN THE CYTOPLASM.
 CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON TYROSINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
 CC GLY-RICH PROTEINS.
 DR HSP, P09651: 10P1.
 DR Interpro: IPR00504; RRM.
 DR PROSITE: PS0102; RRM; PARTIAL.
 DR PROSITE: PS00030; RRM; RNP_1; PARTIAL.
 KM RNA-binding: Ribonucleoprotein; Nuclear protein; phosphorylation;
 KM Methylation.
 FT NON_TER 1 1
 FT NON_CONS 15 16
 FT DOMAIN 32 >48
 FT NON_CONS 48 49
 FT NON_CONS 58 59
 FT NON_CONS 73 74
 FT NON_CONS 83 84
 FT NON_CONS 96 97
 FT DOMAIN 120 195
 FT MOD_RES 119 119
 FT MOD_RES 122 122
 FT MOD_RES 173 173
 FT MOD_RES 192 192
 FT NON_CONS 125 126
 FT NON_TER 195 195
 SO SEQUENCE 195 AA; 20604 MW; CA062DA8520B83DB CRC64;

Query Match 20.9%; Score 177.5; DB 1; Length 195;
 Best Local Similarity 43.4%; Pred. No. 3e-06;
 Matches 46; Conservative 7; Mismatches 30; Indels 23; Gaps 6;
 QY 42 IERAGPRGTYGLTASIGCEGARLRGAGVSSYGYGYRSMGTPYGGYGGY----- 95
 DB 104 IKAATPPDAMMMPPMRGRCGLGLGG-AMVAPGSGFYG-----GGYGGGGYGDADA 154
 QY 96 YG--GYGYDQFGSAYGYGYPYPSGGYGGYGGYGGYGG 139
 DB 155 YGAGGYDYGSGYGGYGG--GYEGY-----GYNGGYGSGPANGG 194

RESULT 14
 ROAL_HUMAN STANDARD; PRT; 371 AA.
 AC P09651;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1 (HELIX-DESTABILIZING
 DE PROTEIN) (SINGLE-STRAND BINDING PROTEIN) (HNRNP CORE PROTEIN A1).
 GN HNRPA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-250 AND 303-371 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89342435; PubMed=2760922;
 RA Blamont G., Buvoil M., Bassi M.T., Morandi C., Cobiachchi F., Riva S.,
 RT "Isolation of an active gene encoding human hnrnp protein A1.
 RT Evidence for alternative splicing."
 RL J. Mol. Biol. 207:491-503(1989).
 RN [2]
 RP SEQUENCE OF 1-250 AND 303-371 FROM N.A.
 RC TISSUE=Fibroblast;
 RX MEDLINE=88233978; PubMed=2836799;

RA Buvoil M., Blamont G., Ghetti A., Riva S., Bassi M.T., Morandi C.;
 RT "DNA cloning of human hnrnp protein A1 reveals the existence of
 RT multiple RNA isoforms."
 RL Nucleic Acids Res. 16:3751-3770(1988).
 RN [3]
 RP SEQUENCE OF 1-250 AND 303-371 FROM N.A.
 RC TISSUE=Lung;
 RA Knudsen S.M., Loeffers H.;
 RT Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 124-250 AND 303-371 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=87053868; PubMed=3023065;
 RA Riva S., Morandi C., Tsoulfas P., Pandolfo M., Blamont G.,
 RA Merrill B., Williams K.R., Maltupa G., Beyreuther K., Werr H.,
 RA Heinrich B., Schaefer K.P.;
 RT "Mammalian single-stranded DNA binding protein UP I is derived from
 RT the hnrnp core protein A1."
 RL EMBO J. 5:2267-2273(1986).
 RN [5]
 RP SEQUENCE OF 251-302 FROM N.A.
 RX MEDLINE=90214633; PubMed=1691095;
 RA Buvoil M., Cobiachchi F., Bestagno M.G., Mangiarotti A., Bassi M.T.,
 RA Blamont G., Riva S.;
 RT "Alternative splicing in the human gene for the core protein A1
 RT generates another hnrnp protein."
 RL EMBO J. 9:1229-1235(1990).
 RN [6]
 RP NUCLEAR LOCALIZATION DOMAIN.
 RX MEDLINE=95247808; PubMed=7730395;
 RA Siomi H., Dreyfuss G.;
 RT "A nuclear localization domain in the hnrnp A1 protein."
 RL J. Cell Biol. 129:551-560(1995).
 RN [7]
 RP NUCLEAR LOCALIZATION DOMAIN, AND NUCLEAR EXPORT.
 RX MEDLINE=96067639; PubMed=8521471;
 RA Michael W.M., Choi M., Dreyfuss G.;
 RT "A nuclear export signal in hnrnp A1: a signal-mediated, temperature-
 RT dependent nuclear protein export pathway."
 RL Cell 83:415-422(1995).
 RN [8]
 RP NUCLEAR LOCALIZATION DOMAIN.
 RX MEDLINE=95286702; PubMed=7769000;
 RA Weighardt F., Blamont G., Riva S.;
 RT "Nucleo-cytoplasmic distribution of human hnrnp proteins: a search
 RT for the targeting domains in hnrnp A1."
 RL J. Cell Sci. 108:545-555(1995).
 RN [9]
 RP 3D-STRUCTURE MODELING OF 106-189.
 RX MEDLINE=91099515; PubMed=2176620;
 RA Ghetti A., Bolognesi M., Cobiachchi F., Morandi C.;
 RT "Modeling by homology of RNA binding domain in A1 hnrnp protein."
 RL FEBS Lett. 277:272-276(1990).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS) OF 8-180.
 RX MEDLINE=97307256; PubMed=9164463;
 RA Shamoo Y., Krueger U., Rice L.M., Williams K.R., Steltz T.A.;
 RT "Crystal structure of the two RNA binding domains of human hnrnp A1
 RT at 1.75-A resolution."
 RL Nat. Struct. Biol. 4:215-222(1997).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 6-181.
 RX MEDLINE=97277240; PubMed=9115444;
 RA Xu R.M., Jokhan L., Cheng X., Mayeda A., Krainer A.R.;
 RT "Crystal structure of human Upl, the domain of hnrnp A1 that contains
 RT two RNA-recognition motifs."
 RL Structure 5:559-570(1997).
 CC -1- FUNCTION: INVOLVED IN THE PACKAGING OF PRE-MRNA INTO HNRNP
 CC PARTICLES, TRANSPORT OF POLY-A MRNA FROM THE NUCLEUS TO THE
 CC CYTOPLASM AND MAY MODULATE SPLICE SITE SELECTION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. SHUTTLES CONTINUOUSLY BETWEEN THE
 CC NUCLEUS AND THE CYTOPLASM ALONG WITH MRNA. COMPONENT OF
 CC RIBONUCLEOSOMES.

```

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A1-A (SHOWN HERE) AND A1-B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. A1-A IS TWENTY TIMES MORE
CC ABUNDANT THAN A1-B.
CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
CC GLY-RICH PROTEINS.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12671; CAA31191.1; -
DR EMBL; X06747; CAA29922.1; ALT_SEQ.
DR EMBL; X04347; CAA27872.1; -
DR EMBL; X79536; CAA56072.1; -
DR PIR; S04617; S04617.
DR PIR; A24894; A24894.
DR PIR; S02061; S02061.
DR PDB; 1HA1; 15-MAY-97.
DR PDB; 1UP1; 17-SEP-97.
DR Aarhus/Ghent-2DPAGE; 207; NEPHGE.
DR Aarhus/Ghent-2DPAGE; 2114; NEPHGE.
DR Aarhus/Ghent-2DPAGE; 3612; NEPHGE.
DR MIM; 164017; -
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 2.
DR Nuclear protein; RNA-binding; Repeat; Ribonucleoprotein;
DR Methylation; Alternative splicing; 3D-structure.
DR INT_MET 0
FT DOMAIN 3 93 GLOBULAR A DOMAIN.
FT DOMAIN 94 184 GLOBULAR B DOMAIN.
FT DOMAIN 13 96 RNA-BINDING (RRM) 1.
FT DOMAIN 104 183 RNA-BINDING (RRM) 2.
FT DOMAIN 217 239 RNA-BINDING RGG-BOX.
FT DOMAIN 194 371 GLY-RICH.
FT DOMAIN 319 356 NUCLEAR TARGETING SEQUENCE (M9).
FT MOD_RES 193 193 METHYLATION (BY SIMILARITY).
FT VAAPPLIC 251 302 MISSING (IN ISOFORM A1-A).
FT MUTAGEN 325 325 G->A: NO NUCLEAR IMPORT NOR EXPORT.
FT MUTAGEN 326 326 G->A: NO NUCLEAR IMPORT NOR EXPORT.
FT MUTAGEN 333 334 GG->LL: NORMAL NUCLEAR IMPORT AND EXPORT.
FT CONFLICT 139 139 R->P (IN REF. 4).
SQ SEQUENCE 371 AA; 38715 MW; B3EFA5A1DB7C26 CRC64;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP-DEPENDENT RNA HELICASE A (NUCLEAR DNA HELICASE II) (NDH II)
DE (DEAD-BOX PROTEIN 9).
CN DDX9 OR NDH2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=9533357; PubMed=7608213;
RA Zhang S., Maacke H., Grosse F.;
RT "Molecular cloning of the gene encoding nuclear DNA helicase II. A
RT bovine homologue of human RNA helicase A and Drosophila Mle
RT protein."
RL J. Biol. Chem. 270:16422-16427(1995).
CC -1- FUNCTION: UNWINDS DOUBLE-STRANDED DNA AND RNA IN A 3' TO 5'
CC DIRECTION. GENERATES MULTIPLE SECONDARY STRUCTURES THAT INFLUENCE
CC RNA-BINDING PROTEINS. MAY PLAY A ROLE IN X-LINKED GENE EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE. DEAD
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 DBRM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82829; CAA58036.1; -
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002464; DEAD_ATP_helcse.
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR001650; Helicase_C.
DR Pfam; PF00035; dsrm; 2.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00358; DSRM; 2.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; 1.
DR PROSITE; PS50137; DS_RBD; 2.
DR Helicase; RNA-binding; DNA-binding; Repeat; Nuclear protein;
KW ATP-binding.
FT DOMAIN 3 71 DBRM 1.
FT DOMAIN 177 249 DBRM 2.
FT NP_BIND 408 415 ATP (BY SIMILARITY).
FT SITE 508 511 DEAD BOX.
FT DOMAIN 583 592 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 1167 1287 ARG/GLY/TYR-RICH.
SQ SEQUENCE 1287 AA; 141943 MW; DC908095AB683EDA CRC64;

```

```

Query Match 20.8%; Score 177; DB 1; Length 371;
Best Local Similarity 33.1%; Pred. No. 5.4e-06;
Matches 47; Conservative 14; Mismatches 53; Indels 28; Gaps 4;

```

```

Query Match 20.7%; Score 176.5; DB 1; Length 1287;
Best Local Similarity 37.9%; Pred. No. 1.6e-05;
Matches 53; Conservative 1; Mismatches 45; Indels 41; Gaps 5;

```

'Thu Jan 17 08:02:54 2002

us-09-554-547-16.rsp

Page 11

Db 1258 YRSGGFORGGGRCYGGY 1277

Search completed: January 16, 2002, 14:58:40
Job time: 192 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 14:55:13 ; Search time 24.43 Seconds
(Without alignments)
922.061 Million cell updates/sec

Title: US-09-554-547-16
Perfect score: 851
Sequence: 1 MKAFFVLSTALTNAR.....GSYGSTYTPNVASAGAAA 154

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_TREMBL_17:*
2: SP_Archaea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207.5	24.4	423	10 Q9FM47	Q9fm47 arabidopsis
2	206.5	24.3	159	5 Q9VE11	Q9ve11 drosophila
3	198	23.3	464	10 Q9FMK8	Q9fmk8 oryza sativ
4	195.5	23.0	245	5 Q9VW59	Q9vw59 drosophila
5	193	22.7	159	11 Q08884	Q08884 mus musculu
6	192	22.6	458	5 Q9N3U1	Q9n3u1 caenorhabdi
7	191.5	22.5	146	5 Q9VFT9	Q9vft9 drosophila
8	191	22.4	141	5 Q9VNN3	Q9vnn3 drosophila
9	190.5	22.4	259	5 Q02049	Q02049 caenorhabdi
10	187	22.0	263	3 Q94056	Q94056 candida alb
11	185.5	21.8	250	5 Q9VX64	Q9vx64 drosophila
12	184.5	21.7	440	5 Q9V6M1	Q9v6m1 drosophila
13	183	21.5	173	5 Q9VX65	Q9vx65 drosophila
14	182.5	21.4	142	5 Q9U517	Q9u517 manduca sex
15	181.5	21.3	305	11 Q9CX86	Q9cx86 mus musculu
16	180	21.2	358	10 Q9LHL2	Q9lhl2 arabidopsis
17	180	21.2	360	5 Q16985	Q16985 araneus dia
18	179.5	21.1	261	10 Q9ZMM2	Q9zmm2 cucumis sat
19	179	21.0	268	10 Q9FI02	Q9fiq2 arabidopsis

20	178.5	21.0	289	10 Q9C909	Q9c909 arabidopsis
21	178.5	21.0	303	5 Q17201	Q17201 bombyx mori
22	177	20.8	691	5 Q9BI03	Q9bi03 dolomedes t
23	176	20.7	196	5 Q24346	Q24346 drosophila
24	175.5	20.6	197	10 Q9ATP3	Q9atp3 pennisetum
25	175.5	20.6	309	10 Q9ENR1	Q9enr1 arabidopsis
26	175	20.6	154	10 Q9SL23	Q9sl23 arabidopsis
27	175	20.6	181	5 Q9VWM4	Q9vwm4 drosophila
28	174.5	20.5	208	10 Q9ZRV2	Q9zrv2 cicier ariet
29	174.5	20.5	988	5 Q17434	Q17434 nephila cla
30	174	20.4	347	10 Q01916	Q01916 phytophthor
31	174	20.4	576	5 Q9VX72	Q9vx72 drosophila
32	173	20.3	193	11 Q70592	Q70592 rattus norv
33	172.5	20.3	255	10 Q9S1H2	Q9s1h2 arabidopsis
34	172	20.2	317	5 Q9W0Z0	Q9w0z0 drosophila
35	172	20.2	404	10 Q9LKA4	Q9lka4 arabidopsis
36	171.5	20.2	165	10 Q24184	Q24184 oryza sativ
37	171.5	20.2	167	10 Q41518	Q41518 triticum ae
38	171.5	20.2	287	5 Q17200	Q17200 bombyx mori
39	171	20.1	126	5 Q44355	Q44355 aequipecten
40	170	20.0	136	5 Q44354	Q44354 aequipecten
41	169.5	19.9	139	5 Q9UAV7	Q9uav7 caenorhabdi
42	169.5	19.9	173	10 Q40052	Q40052 hordeum vul
43	169.5	19.9	2639	5 Q76786	Q76786 antheraea p
44	169	19.9	127	11 P70370	P70370 mus musculu
45	169	19.9	209	5 Q9VX73	Q9vx73 drosophila

ALIGNMENTS

```
RESULT 1
ID Q9FM47 PRELIMINARY: PRT: 423 AA.
AC Q9FM47;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SIMILARITY TO RNA BINDING PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98290546; Pubmed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned pl and tAC clones."
RL DNA Res. 5:41-54(1998).
DR EMBL: AB009052; BAB08520.1; -.
DR InterPro: IPR002952; Eggshell.
DR InterPro: IPR00504; RRM.
DR Pfam: PFM0076; rrm: 2.
DR PRINTS: PR01228; EGGSHELL.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PSS0102; RRM; 2.
DR PROSITE: PSS0030; RRM_RNP_1; 1.
SQ SEQUENCE 423 AA; 42408 MW; DCFB04E24CF10002 CRC64;
```

Query Match 24.4%; Score 207.5; DB 10; Length 423;
Best Local Similarity 40.5%; Pred. No. 6,5e-10;
Matches 60; Conservative 4; Mismatches 51; Indels 33; Gaps 7;

```
OY 23 RLGSGLDFGFGVHNTVAGIERAGRGVGLTASIGCGVAGLGGAGVSSYGVGPS 82
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 220 RFGSRSNFGGCGDYGCGHGGGIGGPGPYKSGGIGGSGSGYGCYGGGFGGCGG 279
```

QY 83 WGYPGYGGY-----GGYGGY--GGYDGFSGSANGYGP-YYGYYY----- 121
 Db 280 YG---GGVGPPEPALCYSGRYGGGGGYNNGYSMGCGGGYGGPGDMYGGSYGEPBG 336
 QY 122 ----PSG-YGGGYG---GSYGGSYGGS 140
 Db 337 GYGPGSGYGGYGGSGIGYGGGMGA 364

RESULT 2

ID Q9VE11 PRELIMINARY: PRT: 159 AA.
 AC Q9VE11;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE EDG91 PROTEIN.
 CN EDG91 OR CG7539.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers V.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abell J.F., Agbayan A., Ah H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Bales R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Beron P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkov D., Botchan M.V., Bouck J., Brokslein P., Brotlier P.,
 Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherly J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbam C.,
 Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Maltel B., McIntosh T.C., Moled M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith H.,
 Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Sheth T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svrystas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zhang X.H., Zheng F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003718; AAF55440.1;
 DR FLYBase: FBgn004554; Edg91.
 DR InterPro: IPR002952; Eggshell.
 DR PRINTS: PRO1228; EGGSHELL.
 SO SEQUENCE 159 AA; 15261 MW; 266216A2A39AA06F CRC64;

Query Match

24.3%; Score 206.5; DB 5; Length 159;

Best Local Similarity 42.0%; Pred. No. 2.6e-10;
 Matches 58; Conservative 7; Mismatches 34; Indels 39; Gaps 10;
 QY 38 LVAGIERAGPRGYPGLTAS-IGGEVGARLGRAGVSS-----YGYGP---SMGYR 87
 Db 14 LIAGGGAAPYKTEESTLGLGGFGSGVLSAGIGVGGGLSGCGGCGYRPGYASGP- 72
 QY 88 GGY-GGYGGYGGYGGYDGFSGAY---GGYGPY-----YGYYPGSG-----YGGYGGGS 132
 Db 73 GYGGGYSGYNGYGS--SGFGGYRPGGYSGFGRPHNHGGYGYGGSYNHGGSGSYGN 130
 QY 133 Y-----GGSYGC 139
 Db 131 YSOSQYSNGYGGGYYGC 148

RESULT 3

ID Q9FWK8 PRELIMINARY: PRT: 464 AA.
 AC Q9FWK8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PUTATIVE RNA BINDING PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriobotryaceae; Oryzae; Oryza.
 CX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Bueli C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 Hsiao J., Zisman V., Pai G., Bowman C.L., Fujii C.Y., Vanken S.E.,
 Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldtlyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNBa007916 genomic sequence.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC026815; AAC21903.1;
 DR InterPro: IPR000504; RRM.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
 SO SEQUENCE 464 AA; 45621 MW; CFC03DD3045E4B92 CRC64;

Query Match 23.3%; Score 198; DB 10; Length 464;
 Best Local Similarity 30.8%; Pred. No. 4.6e-09;
 Matches 60; Conservative 11; Mismatches 48; Indels 76; Gaps 9;

QY 21 AGRGSDLDFTGRHGNLYAGIERAGPRGYPGLTASIGEVG----- 62
 Db 263 AGYGAGYGVGYGHANGTAYVGNVAGGFGVPGYGHYGNPNAPSGYQGCPGRANR 322
 QY 63 ARLGGR--GAVVSSYG--YGPSPW-----GYPGGYG-- 92
 Db 323 GPMVGQAASGYCTGTSYSGNAGYAAAMNNSAGGNAPTSGAAGAGTGYSGYQGYGGYGD 382
 QY 93 ----YGGYGGYGGYDQGG--AYGGYGPYGY-----YPSGY-----GGYGG-- 132
 Db 383 SYGNHGGYGGYGGGDDGAGNPAAGGSGYAGYSGNGSGYPPNAMPDSOGGFGASVN 442
 QY 133 -----YGSYGG 139
 Db 443 GVSQGQSNVSGYGG 457

RESULT 4
 Q9VZ59
 ID Q9VZ59 PRELIMINARY: PRT: 245 AA.
 AC Q9VZ59;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, last annotation update)
 DE CG2157 PROTEIN.
 GN CG2157.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Planck C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkovec D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferrer C., Ferriere S., Fleischmann W.,
 RA Gosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harits N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegum C.,
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Lasko P., Leclay, Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshnell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003484; AAF47968.1; -;
 DR FlyBase, FBgn0030244; CG2157.
 SO SEQUENCE 245 AA; 23982 MW; 838B6D206B40035C CRC64;

Query Match 23.0%; Score 195.5; DB 5; Length 245;
 Best Local Similarity 31.0%; Pred. No. 3.6e-09;
 Matches 61; Conservative 14; Mismatches 67; Indels 55; Gaps 7;

DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, last annotation update)
 DE CG2157 PROTEIN.
 GN CG2157.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Planck C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkovec D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferrer C., Ferriere S., Fleischmann W.,
 RA Gosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harits N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegum C.,
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Lasko P., Leclay, Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshnell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003484; AAF47968.1; -;
 DR FlyBase, FBgn0030244; CG2157.
 SO SEQUENCE 245 AA; 23982 MW; 838B6D206B40035C CRC64;

DB 171 SNYPALSYNSRLGAGA 187

RESULT 5
 ID 008884 PRELIMINARY; PRT; 159 AA.
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HIGH-GLYCINE TYROSINE KERATIN TYPE II.4.
 GN KRTAP6-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKIN;
 RA Aoki N.;
 RT "High-glycine / tyrosine keratin.";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D89902; BAA20281.1; -;
 DR MGD; MGI:1330280; Krtap6-2.
 KW Keratin.
 SO SEQUENCE 159 AA; 15262 MW; 2925AE80A777BB9 CRC64;

Query Match 22.7%; Score 193; DB 11; Length 159;
 Best Local Similarity 46.2%; Pred. No. 3.5e-09;
 Matches 55; Conservative 5; Mismatches 45; Indels 14; Gaps 9;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE HYPOTHEICAL PROTEIN Y47D7A.B.
 GN Y47D7A.B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 MEDLINE=99069613; PubMed=9851916;
 None;
 RT "genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC024790; AAF60630.1; -;
 DR InterPro: IPR002952; Eggshell.
 DR InterPro: IPR000976; Wlms_tumour.
 DR PRINTS; PR01228; EGGSHELL.
 DR PRINTS; PR00049; WILMSTUMOUR.

RESULT 6
 ID 09N301 PRELIMINARY; PRT; 458 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE HYPOTHEICAL PROTEIN Y47D7A.B.
 GN Y47D7A.B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 MEDLINE=99069613; PubMed=9851916;
 None;
 RT "genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC024790; AAF60630.1; -;
 DR InterPro: IPR002952; Eggshell.
 DR InterPro: IPR000976; Wlms_tumour.
 DR PRINTS; PR01228; EGGSHELL.
 DR PRINTS; PR00049; WILMSTUMOUR.

SO SEQUENCE 458 AA; 40183 MW; 3CF054AE11209FE6 CRC64;

Query Match 22.6%; Score 192; DB 5; Length 458;

Best Local Similarity 37.2%; Pred. No. 1.4e-08; Mismatches 59; Indels 24; Gaps 6;

Matches 55; Conservative 10; Mismatches 59; Indels 24; Gaps 6;

QY 22 GRLGSDLDLTFGRVHGNL--VAGIERAPRGYPGLTASIGGEVGARLGGAGVGVSSYSG 79

Db 227 GYAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 286

QY 80 YPSMCKYPPGGYGGYGGY-----GGYGGYDQ---GFGSAYGGYPPGYGYGYSSGYG 128

Db 287 ---GYAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 337

QY 129 YGGSYGS--YGSYTPNVASAGAAA 154

Db 338 FGGSGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 365

RESULT 7

QYVFT9 PRELIMINARY; PRT; 146 AA.

AC QYVFT9; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE CG9269 PROTEIN.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

SEQUENCE FROM N.A.

STRAIN-BERKELEY;

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,

RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster";

RL Science 287:2185-2195(2000).

DR EMBL: AE003701; AAF54960.1; -

DR FlyBase: FBgn0038161; CG9269.

SO SEQUENCE 146 AA; 1478 MW; 35DD14C33544BFA3 CRC64;

Query Match 22.5%; Score 191.5; DB 5; Length 146;

Best Local Similarity 50.0%; Pred. No. 4.3e-09; Mismatches 26; Indels 17; Gaps 7;

Matches 46; Conservative 3; Mismatches 26; Indels 17; Gaps 7;

QY 65 LGRAGVSSYGYGYPGSGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGG 124

Db 46 LGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGG 98

QY 125 --YGS--GYGSGYGG--SYGSGY--YPNV 146

Db 99 NLGGSGGLAGAGYGRPYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGG 130

RESULT 8

QYVNN3 PRELIMINARY; PRT; 141 AA.

AC QYVNN3; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE CG15597 PROTEIN.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

SEQUENCE FROM N.A.

STRAIN-BERKELEY;

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,

RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CG10597 PROTEIN.
GN CG10597.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cline S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster".
RL Science 287:2185-2195(2000).
DR EMBL: AEO03504; AAF48714.1; -
DR FlyBase: FBgn0030832; CG10597.
DR InterPro: IPR002952; Eggshell.
DR PRINTS: PR01228; EGGSHLL.
SQ SEQUENCE 250 AA; 23045 MW; 9A7DE402ED0F03BE CRC64;

Query Match 21.8%; Score 185.5; DB 5; Length 250;
Best Local Similarity 33.7%; Pred. No. 2.5e-08;
Matches 66; Conservative 12; Mismatches 59; Indels 59; gaps 8;

QY 9 LLSAALTAALMA--RAGRLGSDLDTFGRVHG-----NLVAG 41
DB 57 LLSAALTAALMAGGLGAGAGGCVLPGGGRRGGGCGRPAICAGLVSHVSOGATKTHIG 116
QY 42 IERAGPRG-YPGLTASIGEGYARLGRAGY----GVSSYGYGYPSKCYPGYGYG-GYGG 95
DB 117 GSGAGGAAVGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 176
QY 96 YGAGYGGYDGGGSAVGGYGPY-----YGYRPSGYYGGYGGGSGYGG 135

DB 177 AGGAGGW-OGAGAGRGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAYG 235
QY 136 SYGSGYTPNPARASG 151
DB 236 AAGG-----GSRGAG 246

RESULT 12
ID 09V6M1 PRELIMINARY; PRT; 440 AA.
AC 09V6M1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CG4663 PROTEIN.
GN CG4663.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cline S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster".
RL Science 287:2185-2195(2000).
DR EMBL: AEO03819; AAF58402.1; -
DR FlyBase: FBgn0033812; CG4663.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PSS0002; SH3; 1.
SQ SEQUENCE 440 AA; 46660 MW; 3ABEC719DDE9C50D CRC64;

Query Match 21.7%; Score 184.5; DB 5; Length 440;

QY	35	HCNLVAGIERACPRCTPCL-----TASIGEVGARLGRAGVSVSYGYCPSPMKXPY	87
DB	44	YGNVRA---LPCPAQPPPLPQSPFOOTQOYGC-----FGSYGCNNYGL	84
QY	88	GGYGVG----CYGCGTCGDOGFSAGYAGCPYCYGYTSYGGGYSYGSY	141
DB	85	GGYGGFNSGAFYGGGLGFGSGLGSGFGS---GY---GGYGGYGGGCGGCGY	133
RESULT	13		
Q9VX65		PRELIMINARY:	173 AA.
AC	Q9VX65		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	ANON-PEN15EF PROTEIN.		
OS	CG10598 OR ANON-PEN15EF.		
OC	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY:		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Mortman J.R., Yendell M.D., Zheng Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abrial J.F., Aydayani A., An H.-J., Andrews-Pfankuch C., Balowin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mottel B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,		
RA	Palazzolo M., Plutman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington C., Saunders R.D.C., Scheller F., Shen H.,		
RA	Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287:2185-2195(2000).		
DR	EMBL: AE003504; AAF48713.1; -		
DR	PIRBASE: FBGN0003057; CG10598.		
DR	InterPro: IPR002952; Eggshell.		
DR	PRINTS: PR01228; EGGSHLL.		
SO	SEQUENCE 173 AA; 15408 MW; 4A337106E4FA1F25 CRC64;		

```

Query Match 21.5%; Score 183; DB 5; Length 173;
Best Local Similarity 34.1%; Pred. No 2.7e-08;
Matches 56; Conservative 13; Mismatches 71; Indels 24; Gaps 6;

OY 1 MKAFVLLSTLAULT--NAARARIGSLDLPFRGHNLVAGIERAGPRGPGTLASIG 58
      | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 1 MKPFACILVIAAILVSGASIGLLGGGGGGGGIGGA--GGIGQLQSKJLGGG 59
      | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 59 GEVVARLGRAG--VEVSSYGYGYPSPWGPY--GGYGGYGGYGGYGGYDQFGAAG---- 111
      | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 60 GGLGGLSGGLGKLGKGGGGGGG--SGGYSNNGGYSGGGGYSGGGYSGGGYSGGYAAPR 118
      | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 112 -----GYPGYGYPYPSPGYGGGYSGGYSGGYSYT 142
      | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 119 PVEKVIYKVINEGYSGGYSAGNAGSGGYSGGYSAGNAGSGGYS 162

RESULT 14
O9U517 PRELIMINARY: PRT: 142 AA.
AC O9U517:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE PUTATIVE CUTICLE PROTEIN.
OS Manduca sexta (tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dityrista;
OC SpHINGLODEA; SpHINGIDAE; SpHINGINAE; Manduca.
OX NCBI_TaxID=7130;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20099029; PubMed=10620045;
RA Robertson H.M., Marcos R., Sears C.R., Todres E.Z., Walden K.K.O.,
RA Nardel J.B.;
RT "Diversity of odorant binding proteins revealed by an expressed
RT sequence tag project on male Manduca sexta moth antennae.";
RL Insect Mol. Biol. 8:501-518 (1999).
DR EMBL; AF117571; AAF16693.1; --
SO SEQUENCE 142 AA; 14112 MW; 17230DEB564DE0 CXC64;

Query Match 21.4%; Score 182.5; DB 5; Length 142;
Best Local Similarity 43.6%; Pred. No. 2.4e-08;
Matches 51; Conservative 4; Mismatches 37; Indels 25; Gaps 7;

OY 39 YAGIERAGGR-GYPLTISIGGEVAGARLGRAGVYSSGYGYPSPWGYGGYGGY---G 94
      | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 32 YGGLYAGAGVGYDGL--GYGGYGLGYSGPSVAVSKVATTHAGGLAGYGGYGNLGL 89
      | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 95 GYGYGYGDOGFG-----SAYGGYGYGYPYPSPGY--GGYGGYSGGSYG 138
      | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 90 GYSGLGAGVGLGHGYSAAPVAYSNQAAVAGYGGY---GHGLGAGGV--GYYG 140
      | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 15
O9CX86 PRELIMINARY: PRT: 305 AA.
AC O9CX86:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 3010025E17RIK.
DE 3010025E17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;

```


**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)